

A forward-adjoint operator pair based on the elastic wave equation for use in transcranial photoacoustic computed tomography

Kenji Mitsuhashi , Joemini Poudel , Thomas P. Matthews , Alejandro Garcia-Urbe ,
Lihong V. Wang ^{*}, and Mark A. Anastasio [†]

Abstract.

Photoacoustic computed tomography (PACT) is an emerging imaging modality that exploits optical contrast and ultrasonic detection principles to form images of the photoacoustically induced initial pressure distribution within tissue. The PACT reconstruction problem corresponds to an inverse source problem in which the initial pressure distribution is recovered from measurements of the radiated wavefield. A major challenge in transcranial PACT brain imaging is compensation for aberrations in the measured data due to the presence of the skull. Ultrasonic waves undergo absorption, scattering and longitudinal-to-shear wave mode conversion as they propagate through the skull. To properly account for these effects, a wave-equation-based inversion method should be employed that can model the heterogeneous elastic properties of the skull. In this work, a forward model based on a finite-difference time-domain discretization of the three-dimensional elastic wave equation is established and a procedure for computing the corresponding adjoint of the forward operator is presented. Massively parallel implementations of these operators employing multiple graphics processing units (GPUs) are also developed. The developed numerical framework is validated and investigated in computer-simulation and experimental phantom studies whose designs are motivated by transcranial PACT applications.

Key words. Photoacoustic computed tomography, transcranial imaging, image reconstruction, elastic wave equation

1. Introduction. Photoacoustic computed tomography (PACT) is a noninvasive imaging modality that exploits the optical absorption contrast of tissue with the high spatial resolution of ultrasound imaging techniques [42, 27, 55]. In PACT, the target is illuminated with a short optical pulse that results in the generation of acoustic pressure signals via the thermoacoustic effect [53, 40]. The propagated acoustic pressure signals are then detected by use of a collection of wideband ultrasonic transducers that are located outside the support of the object. Typically, the measured pressure signals are employed to estimate the induced initial pressure distribution or, equivalently, if the Grüneisen parameter is known, the absorbed optical energy distribution.

Transcranial brain imaging represents an important application that may benefit significantly by the development of PACT methods. Existing human brain imaging modalities include X-ray computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and ultrasonography. However, all these modalities suffer from significant shortcomings. X-ray CT, PET, and MRI are expensive and employ bulky and generally non-portable imaging equipment. Moreover, X-ray CT and PET employ ionizing radiation and are therefore not suitable for longitudinal studies and MRI-based methods are generally slow. Ultrasonography is an established portable pediatric brain imaging modality

^{*}Department of Biomedical Engineering, Washington University in St.Louis, St.Louis, MO USA (lh-wang@wustl.edu).

[†]Department of Biomedical Engineering, Washington University in St.Louis, St.Louis, MO USA (anastasio@wustl.edu).

that can operate in near real-time, but its image quality degrades severely when employed after the closure of the fontanelles. On the other hand, PACT can be implemented in near real-time, does not employ ionizing radiation, is much less costly than either MRI, PET or X-ray CT, and can provide both anatomical and functional information.

In vivo transcranial PACT studies have revealed structure and hemodynamic responses in small animals [50, 30, 55]. In these small animal studies, PACT was used to visualize brain structure, brain lesions and neurofunctional activities such as cerebral hemodynamic responses to hyperoxia and hypoxia, and cerebral cortical responses to various forms of stimulation [50, 30]. In addition to providing functional and structural information about the brain, transcranial PACT has also been utilized to study the formation of cerebral edema and its expansion and recovery [55]. Hence, transcranial PACT is a neuroimaging modality that holds promise for applications in neurophysiology, neuropathology and neurotherapy. Because the skulls in these small animal studies were relatively thin (~ 1 mm), they did not significantly aberrate the photoacoustic wavefields. As such, conventional backprojection (BP) methods that ignored the presence of the skull and assumed a homogeneous lossless fluid medium were employed for image reconstruction with good success. However, PA signals can be significantly aberrated by thicker skulls present in adolescent and adult primates. To render PACT an effective imaging modality for use with transcranial imaging in large primates, it is necessary to develop image reconstruction methodologies that can accurately compensate for skull-induced aberrations of the recorded PA signals.

Towards this goal, ex vivo studies involving primate heads have also been conducted [54, 24, 55, 56, 39, 22]. In such applications, the effects of the skull on the recorded photoacoustic wavefield were no longer negligible. To address this, a subject-specific imaging model that approximately describes the interaction of the photoacoustic wavefield with the skull was developed [22]. This imaging model was established by use of adjunct imaging data, such as X-ray CT, that specified the skull morphology and composition [33, 25, 2]. It was demonstrated that image reconstruction based on the subject-specific imaging model yielded images that contained significantly reduced artifact levels compared to those reconstructed by use of a conventional BP method [22]. However, a limitation of that work was that it assumed a fluid medium and therefore assumed a simplified wave propagation model in which longitudinal-to-shear-wave mode conversion within the skull [44, 51, 20], which is an elastic solid, was neglected. The deleterious effects of making such an approximation in transcranial PACT have been studied previously [45].

To circumvent limitations of previous approaches, in this work a numerical framework for image reconstruction in transcranial PACT based on an elastic wave equation that describes an linear isotropic, lossy and heterogeneous medium is developed and investigated. Similar to the work by Huang *et al.* [22], estimates of the acoustic parameters of the skull are assumed to be obtainable from adjunct image data. In transcranial ultrasound therapy applications, the skull's acoustic parameters are routinely estimated in this way [33, 25, 2, 34]. However, unlike previous methods, the skull is treated as an elastic solid and therefore longitudinal-to-shear-wave mode conversion within the skull is modeled.

The primary contributions of this work are the establishment of a discrete forward operator (i.e., imaging model) for transcranial PACT that is based on the three-dimensional (3D) elastic wave equation and a procedure to implement an associated matched adjoint operator.

Specifically, the finite-difference time-domain method (FDTD) is adopted for implementing the forward operator. Both the forward and adjoint operators are implemented using multiple graphics processing units (GPUs). In certain cases, the adjoint operator may serve as a useful image reconstruction operator. More generally, however, the ability to compute the adjoint operator will permit application of gradient-based iterative reconstruction algorithms that seek to minimize a specified objective function. The developed numerical framework is validated and investigated in computer-simulation and experimental phantom studies.

The paper is organized as follows. In Section 2, the salient imaging physics and image reconstruction principles are reviewed briefly. The explicit formulation of the forward operator is described in Section 3, and the implementation of the corresponding discrete adjoint operator is described in Section 4. The forward operator is validated in Section 5, in which simulated and analytically-produced measurement data are compared. Finally, the forward and adjoint operators are applied in image reconstruction studies involving computer-simulated and experimental data in Sections 6 and 7, respectively.

2. Background. The principles of photoacoustic wavefield generation and propagation in an elastic medium are described below in their continuous and discrete forms. The discrete description is based on the FDTD method [8, 37, 38, 48]. The FDTD method is described by use of a matrix notation, which subsequently will facilitate the computation of the matched adjoint operator.

2.1. Photoacoustic wavefield propagation: Continuous formulation. Let the photoacoustically-induced stress tensor at location $\mathbf{r} \in \mathbb{R}^3$ and time $t \geq 0$ be defined as

$$(1) \quad \boldsymbol{\sigma}(\mathbf{r}, t) \equiv \begin{bmatrix} \sigma^{11}(\mathbf{r}, t) & \sigma^{12}(\mathbf{r}, t) & \sigma^{13}(\mathbf{r}, t) \\ \sigma^{21}(\mathbf{r}, t) & \sigma^{22}(\mathbf{r}, t) & \sigma^{23}(\mathbf{r}, t) \\ \sigma^{31}(\mathbf{r}, t) & \sigma^{32}(\mathbf{r}, t) & \sigma^{33}(\mathbf{r}, t) \end{bmatrix},$$

where $\sigma^{ij}(\mathbf{r}, t)$ represents the stress in the i^{th} direction acting on a plane perpendicular to the j^{th} direction. Additionally, let $p_0(\mathbf{r})$ denote the photoacoustically-induced initial pressure distribution within the object, and $\dot{\mathbf{u}}(\mathbf{r}, t) \equiv (\dot{u}^1(\mathbf{r}, t), \dot{u}^2(\mathbf{r}, t), \dot{u}^3(\mathbf{r}, t))$ represent the vector-valued acoustic particle velocity. Let $\rho(\mathbf{r})$ denote medium's density distribution and $\lambda(\mathbf{r})$, $\mu(\mathbf{r})$ represent the Lamé parameters that describe the full elastic tensor of the linear isotropic media. All functions in this work are assumed to be bounded and compactly supported.

The compressional and shear wave propagation speeds are given by,

$$(2a) \quad c_l(\mathbf{r}) = \sqrt{\frac{\lambda(\mathbf{r}) + 2\mu(\mathbf{r})}{\rho(\mathbf{r})}} \text{ and } c_s(\mathbf{r}) = \sqrt{\frac{\mu(\mathbf{r})}{\rho(\mathbf{r})}},$$

respectively. In transcranial PACT imaging applications, the acoustic absorption is not negligible. Here, acoustic absorption within the skull is described by a diffusive absorption model [41]. The diffusive absorption model ignores the fact that the wavefield absorption is dependent on temporal frequency. This model, however, is reasonable for cases where the bandwidth of the photoacoustic signals are limited. Moreover, the model also assumes that the shear absorption to compressional absorption ratio is given by the compressional velocity to shear

velocity ratio. This is approximately true in bone, because the slower shear waves are, in fact, more attenuated than the faster compressive waves [41].

In a 3D heterogeneous linear isotropic elastic medium with an acoustic absorption coefficient $\alpha(\mathbf{r})$, the propagation of $\dot{\mathbf{u}}(\mathbf{r}, t)$ and $\boldsymbol{\sigma}(\mathbf{r}, t)$ can be modeled by the following two coupled equations [8, 48, 32, 1]:

$$(3a) \quad \partial_t \dot{\mathbf{u}}(\mathbf{r}, t) + \alpha(\mathbf{r}) \dot{\mathbf{u}}(\mathbf{r}, t) = \frac{1}{\rho(\mathbf{r})} (\nabla \cdot \boldsymbol{\sigma}(\mathbf{r}, t))$$

and

$$(3b) \quad \partial_t \boldsymbol{\sigma}(\mathbf{r}, t) = \lambda(\mathbf{r}) \mathbf{tr}(\nabla \dot{\mathbf{u}}(\mathbf{r}, t)) \mathbf{I} + \mu(\mathbf{r}) (\nabla \dot{\mathbf{u}}(\mathbf{r}, t) + \nabla \dot{\mathbf{u}}(\mathbf{r}, t)^T),$$

subject to the initial conditions

$$(3c) \quad \boldsymbol{\sigma}_0(\mathbf{r}) \equiv \boldsymbol{\sigma}(\mathbf{r}, t)|_{t=0} = -\frac{1}{3} p_0(\mathbf{r}) \mathbf{I}, \quad \dot{\mathbf{u}}(\mathbf{r}, t)|_{t=0} = 0.$$

Here, $\mathbf{tr}(\cdot)$ is the operator that calculates the trace of a matrix and $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ is the identity matrix. In Eqn. (3c), it has been assumed that the object function $p_0(\mathbf{r})$ is compactly supported in a fluid medium where the shear modulus $\mu(\mathbf{r}) = 0$. In transcranial PACT, this corresponds to the situation where the initial photoacoustic wavefield is produced within the soft tissue enclosed by the skull. Note, the initial conditions defined in Eqn. (3) would have to be modified if the photoacoustically-induced stress is generated within the elastic media.

2.2. Photoacoustic wavefield propagation: Discrete formulation. The FDTD method is employed to propagate the photoacoustic wavefield forward in space and time by computing numerical solutions to the coupled equations given in Eqn. (3) [37]. In the FDTD method, at a given temporal step, each grid point is updated based on the local information around that same point. As the FDTD method utilizes local information, it lends itself to distributed programming across multiple devices. Since the latency in communication between devices is a limiting factor on computational efficiency, spectral and pseudo-spectral methods [10, 46, 19, 47] that use global information at each temporal update are not as efficient for distributed programming. In addition, because of the simplicity of the FDTD method to model elastic wave propagation, it still remains very widely employed in seismology [37].

The salient features of the FDTD method that will underlie the discrete PACT imaging model are described below. In the initial applications of the FDTD method to the elastic wave equation, all functions (e.g., stress tensor and particle velocity) were sampled at the same grid positions [1, 8]. However, such conventional grid schemes were found to possess limitations. Problems with grid dispersion and instabilities in media possessing high Poisson's ratios led Virieux [48, 49] to introduce the staggered-grid velocity-stress finite difference schemes for modeling elastic wave propagation. In our study we will employ the introduced the 4th-order staggered-grid FD scheme to compute the numerical spatial derivatives. The 4th-order staggered-grid scheme in 3D has been shown to reduce the computer memory requirements by at least eight times compared to 2nd-order schemes without any loss in accuracy [37]. A staggered-grid FD cell with positions where the particle-velocity components, stress-tensor components, density, elastic material parameters are sampled is illustrated in Figure 1 [37].

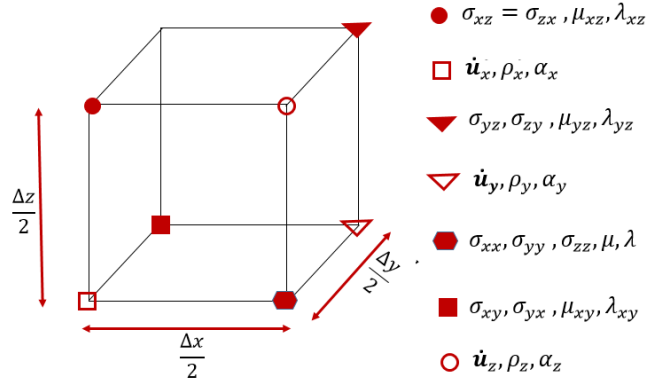


Figure 1. A staggered-grid FD cell with positions of the wavefield variables [37].

Note that for the FD scheme, the material properties, stress, and particle velocity functions, are sampled at different points of the staggered FD cell depicted in Figure 1. Let the set of position vectors $\{\mathbf{r}_1^i, \mathbf{r}_2^i, \dots, \mathbf{r}_N^i \in \mathbb{R}^3\}$ specify the locations where u^i , for $i = 1, 2, 3$, is sampled and $\{\mathbf{r}_1^{jk}, \mathbf{r}_2^{jk}, \dots, \mathbf{r}_N^{jk} \in \mathbb{R}^3\}$ specify the locations where σ^{jk} (for $j, k = 1, 2, 3$) is sampled. Here, $N = N_1 N_2 N_3$ specifies the number of vertices of a 3D Cartesian grid, where N_i denotes the number of vertices along the i^{th} direction. Additionally, let $m\Delta t$, $m \in \mathbb{Z}^*$, $\Delta t \in \mathbb{R}^+$, denote discretized values of the temporal coordinate t , where \mathbb{Z}^* and \mathbb{R}^+ denote the sets of non-negative integers and positive real numbers, respectively. Lexicographically ordered vector representations of the components of the sampled particle velocity and stress tensor will be denoted as

$$(4a) \quad \dot{\mathbf{u}}_m^i = [\dot{\mathbf{u}}^i(\mathbf{r}_1^i, m\Delta t), \dots, \dot{\mathbf{u}}^i(\mathbf{r}_N^i, m\Delta t)]^T$$

and

$$(4b) \quad \boldsymbol{\sigma}_m^{jk} = [\sigma^{jk}(\mathbf{r}_1^{jk}, m\Delta t), \dots, \sigma^{jk}(\mathbf{r}_N^{jk}, m\Delta t)]^T.$$

Let \mathbf{Q}_i , \mathbf{A}_i , $\mathbf{\Lambda}_{jk}$ and $\mathbf{M}_{jk} \in \mathbb{R}^{N \times N}$ be matrices describing the elastic properties of the medium defined as

$$(5a) \quad \mathbf{Q}_i \equiv \text{diag} [\rho(\mathbf{r}_1^i), \dots, \rho(\mathbf{r}_N^i)],$$

$$(5b) \quad \mathbf{A}_i \equiv \text{diag} [\alpha(\mathbf{r}_1^i), \dots, \alpha(\mathbf{r}_N^i)],$$

$$(5c) \quad \mathbf{\Lambda}_{jk} \equiv \text{diag} [\lambda(\mathbf{r}_1^{jk}), \dots, \lambda(\mathbf{r}_N^{jk})],$$

and

$$(5d) \quad \mathbf{M}_{jk} \equiv \text{diag} [\mu(\mathbf{r}_1^{jk}), \dots, \mu(\mathbf{r}_N^{jk})],$$

where $\text{diag}(a_1, \dots, a_N)$ denotes a diagonal matrix whose diagonal entries starting in the upper left corner are a_1, \dots, a_N . Let $\bar{\mathbf{u}}_{m-\frac{1}{2}} \in \mathbb{R}^{3N \times 1}$ and $\bar{\boldsymbol{\sigma}}_m \in \mathbb{R}^{6N \times 1}$ be the concatenation of the

unique components of the particle velocity and stress tensor, defined as

$$(6a) \quad \bar{\mathbf{u}}_{m-\frac{1}{2}} \equiv \left[\dot{\mathbf{u}}_{m-\frac{1}{2}}^1, \dot{\mathbf{u}}_{m-\frac{1}{2}}^2, \dot{\mathbf{u}}_{m-\frac{1}{2}}^3 \right]^T$$

and

$$(6b) \quad \bar{\boldsymbol{\sigma}}_m \equiv [\boldsymbol{\sigma}_m^{11}, \boldsymbol{\sigma}_m^{22}, \boldsymbol{\sigma}_m^{33}, \boldsymbol{\sigma}_m^{23}, \boldsymbol{\sigma}_m^{13}, \boldsymbol{\sigma}_m^{12}]^T.$$

Note that because of the symmetry of the stress tensor (e.g., $\boldsymbol{\sigma}^{ij} = \boldsymbol{\sigma}^{ji}$), it is not necessary to calculate all nine components of the second-order tensor. Here, we have chosen the ordering of $\bar{\boldsymbol{\sigma}}_m$ to follow Voigt notation [21].

To simplify the subsequent presentation, the following operators are defined:

$$(7a) \quad \mathbf{J}_i \bar{\mathbf{u}}_{m-\frac{1}{2}} \equiv (\mathbf{I}_{N \times N} - \Delta t \mathbf{A}_i) \dot{\mathbf{u}}_{m-\frac{1}{2}}^i$$

$$(7b) \quad \Phi_i \bar{\boldsymbol{\sigma}}_m \equiv \Delta t \mathbf{Q}_i^{-1} \sum_{j=1}^3 \partial_j \boldsymbol{\sigma}_m^{ij}$$

$$(7c) \quad \Psi_{ij} \bar{\mathbf{u}}_{m-\frac{1}{2}} \equiv \Delta t \left[\delta_{ij} \Lambda_{ij} \sum_{k=1}^3 \partial_k \dot{\mathbf{u}}_{m-\frac{1}{2}}^k + \mathcal{M}_{ij} \left(\partial_i \dot{\mathbf{u}}_{m-\frac{1}{2}}^j + \partial_j \dot{\mathbf{u}}_{m-\frac{1}{2}}^i \right) \right],$$

where δ_{ij} is the Kronecker delta, $\mathbf{J}_i, \Psi_{ij} \in \mathbb{R}^{N \times 3N}$, $\Phi_i \in \mathbb{R}^{N \times 6N}$, and ∂_i denotes the partial derivative with respect to the i^{th} spatial coordinate. These operators will allow us to compactly express the discrete form of Eqn. (3). In addition, the spatial derivatives defined in Eqn. (7) are calculated using a 4th-order finite-difference scheme. The 4th-order finite-difference approximation to the first derivative along any arbitrary direction x in a staggered grid setup is given by [37]

$$(8) \quad \frac{d\phi}{dx}(x_0) = \frac{1}{\Delta x} \left[-\frac{1}{24} \left(\phi(x_0 + \frac{3}{2}\Delta x) - \phi(x_0 - \frac{3}{2}\Delta x) \right) + \frac{9}{8} \left(\phi(x_0 + \frac{1}{2}\Delta x) - \phi(x_0 - \frac{1}{2}\Delta x) \right) \right],$$

where $\phi(x)$ is an arbitrary field variable (e.g., stress tensor or particle velocity for elastic wave equation). Additionally, define the operators

$$(9a) \quad \mathbf{J} \equiv [\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3]^T,$$

$$(9b) \quad \Phi \equiv [\Phi_1, \Phi_2, \Phi_3]^T, \text{ and}$$

$$(9c) \quad \Psi \equiv [\Psi_{11}, \Psi_{22}, \Psi_{33}, \Psi_{23}, \Psi_{13}, \Psi_{12}]^T,$$

where $\mathbf{J} \in \mathbb{R}^{3N \times 3N}$, and $\Phi \in \mathbb{R}^{3N \times 6N}$, $\Psi \in \mathbb{R}^{6N \times 3N}$. In terms of these quantities, the discretized forms of Eqns. (3a) and (3b) can be expressed as

$$(10a) \quad \bar{\mathbf{u}}_{m+\frac{1}{2}} = \mathbf{J} \bar{\mathbf{u}}_{m-\frac{1}{2}} + \Phi \bar{\boldsymbol{\sigma}}_m$$

$$(10b) \quad \bar{\boldsymbol{\sigma}}_{m+1} = \bar{\boldsymbol{\sigma}}_m + \Psi \bar{\mathbf{u}}_{m+\frac{1}{2}}.$$

2.3. Specification of the image reconstruction problem. The PACT image reconstruction problem addressed in this work is to obtain an estimate of the photoacoustically induced initial pressure distribution $p_0(\mathbf{r})$ from pressure measurements recorded by a collection of ultrasonic transducers surrounding the object. Let $\hat{\mathbf{p}}_m \equiv (p(\mathbf{r}_0^d, m\Delta t), \dots, p(\mathbf{r}_{L-1}^d, m\Delta t))^T$ denote the measured pressure wavefield at time $t = m\Delta t$, for $m = 0, \dots, M-1$, where M is the total number of time steps, and let $\mathbf{r}_l^d \in \mathbb{R}^3$, for $l = 0, \dots, L-1$ denote the positions of the L ultrasonic transducers that reside outside the support of the object $p_0(\mathbf{r})$. Here, for simplicity, we neglect the acousto-electrical impulse response of the ultrasonic transducers and assume each transducer is point-like. However, we can incorporate the spatial and electrical impulse responses of the transducers in the developed discrete imaging model [23]. In addition, the acoustic parameters of the medium are assumed to be known.

A general form of the discrete PACT imaging model can be expressed as

$$(11) \quad \hat{\mathbf{p}} = \mathbb{H}\mathbf{p}_0,$$

where the $LM \times 1$ vector

$$(12) \quad \hat{\mathbf{p}} \equiv \begin{bmatrix} \hat{\mathbf{p}}_0 \\ \hat{\mathbf{p}}_1 \\ \vdots \\ \hat{\mathbf{p}}_{M-1} \end{bmatrix}$$

represents the measured pressure data corresponding to all transducer locations and temporal samples. Additionally, $\mathbf{p}_0 \in \mathbb{R}^{N \times 1}$ is the discrete representation of the sought after initial pressure distribution within the object that is given by

$$(13) \quad \mathbf{p}_0 = - \sum_i \sigma_0^{ii},$$

where σ_0^{jk} is the photoacoustically induced initial stress distribution (Eqn.(4b) with $m = 0$). The $LM \times N$ matrix \mathbb{H} represents the discrete imaging operator (that specifies the forward model), also referred to as the system matrix. The construction of the system matrix \mathbb{H} is based on the initial value problem defined in Eqn. (3), and is described in great detail in Section 3.

The image reconstruction task in a discrete setting is to determine an estimate of \mathbf{p}_0 from knowledge of the measured data $\hat{\mathbf{p}}$. Multiple classes of iterative image reconstruction algorithms require the actions of the operators \mathbb{H} and its adjoint \mathbb{H}^\dagger to be computed repeatedly [9, 13, 3]. Moreover, in some cases, the adjoint operator may serve as a useful heuristic reconstruction operator. Methods for implementing these operators are described below.

3. Explicit formulation of discrete imaging model. The FDTD method for numerically solving the photoacoustic wave equation in elastic linear isotropic media described in Section 2.2 will be employed to implement the action of the system matrix \mathbb{H} . In this section, we provide an explicit representation of \mathbb{H} that will subsequently be employed to determine \mathbb{H}^\dagger . Equation (10) can be described by a single matrix equation to determine the updated wavefield variables after a time step Δt as

$$(14) \quad \mathbf{v}_{m+1} = \mathbf{W}\mathbf{v}_m,$$

where

$$(15) \quad \mathbf{v}_m = \begin{bmatrix} \bar{\mathbf{u}}_{m-\frac{1}{2}} \\ \bar{\boldsymbol{\sigma}}_m \end{bmatrix},$$

and $\mathbf{W} \in \mathbb{R}^{9N \times 9N}$ is the propagator matrix defined as

$$(16) \quad \mathbf{W} \equiv \begin{bmatrix} \mathbf{J} & \boldsymbol{\Phi} \\ \boldsymbol{\Psi}\mathbf{J} & \mathbf{I}_{6N \times 6N} + \boldsymbol{\Psi}\boldsymbol{\Phi} \end{bmatrix}.$$

The wavefield quantities can be propagated forward in time from $t = 0$ to $t = (M - 1)\Delta t$ as

$$(17) \quad \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_{M-1} \end{bmatrix} = \mathbf{T}_{M-1} \cdots \mathbf{T}_1 \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{0}_{9N \times 1} \\ \vdots \\ \mathbf{0}_{9N \times 1} \end{bmatrix},$$

where the $9NM \times 9NM$ matrices \mathbf{T}_m ($m = 1, \dots, M - 1$) are defined in terms of \mathbf{W} as

$$(18) \quad \mathbf{T}_m \equiv \begin{bmatrix} \mathbf{I}_{9N \times 9N} & \cdots & \mathbf{0}_{9N \times 9N} \\ \vdots & \ddots & \vdots & \mathbf{0}_{(m+1) \cdot 9N \times (M-m) \cdot 9N} \\ \mathbf{0}_{9N \times 9N} & \cdots & \mathbf{I}_{9N \times 9N} \\ \mathbf{0}_{9N \times 9N} & \cdots & \mathbf{W} \\ \mathbf{0}_{(M-m-1) \cdot 9N \times m \cdot 9N} & & & \mathbf{0}_{(M-m-1) \cdot 9N \times (M-m) \cdot 9N} \end{bmatrix}$$

with \mathbf{W} residing between the $(9N(m - 1) + 1)^{th}$ to $9Nm^{th}$ columns and the $(9Nm + 1)^{th}$ to $9N(m + 1)^{th}$ rows of \mathbf{T}_m .

From the initial conditions in Eqn. (3c), the vector $(\mathbf{v}_0, \mathbf{0}_{9N \times 1}, \dots, \mathbf{0}_{9N \times 1})^T$ can be computed from the initial pressure distribution \mathbf{p}_0 as

$$(19) \quad \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{0}_{9N \times 1} \\ \vdots \\ \mathbf{0}_{9N \times 1} \end{bmatrix} = \mathbf{T}_0 \mathbf{p}_0,$$

where

$$(20) \quad \mathbf{T}_0 \equiv [\boldsymbol{\tau}, \mathbf{0}_{9N \times N}, \dots, \mathbf{0}_{9N \times N}]^T \in \mathbb{R}^{9NM \times N}$$

and

$$(21) \quad \boldsymbol{\tau} \equiv [\mathbf{0}_{3N \times N}, -\frac{1}{3}\mathbf{I}_{N \times N}, -\frac{1}{3}\mathbf{I}_{N \times N}, -\frac{1}{3}\mathbf{I}_{N \times N}, \mathbf{0}_{3N \times N}]^T \in \mathbb{R}^{9N \times N},$$

with \mathbf{p}_0 being specified by Eqn. (13).

In general, the transducer locations \mathbf{r}_l^d at which the PA data $\hat{\mathbf{p}}$ are recorded will not coincide with the vertices of the 3-D Cartesian grid at which the propagated field quantities

are computed. The measured data $\hat{\mathbf{p}}$ can be related to the computed field quantities via an interpolation operation defined as

$$(22) \quad \hat{\mathbf{p}} = \mathbf{M} \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_{M-1} \end{bmatrix}, \text{ where } \mathbf{M} \equiv \begin{bmatrix} \mathbf{\Theta} & \mathbf{0}_{L \times 9N} & \cdots & \mathbf{0}_{L \times 9N} \\ \mathbf{0}_{L \times 9N} & \mathbf{\Theta} & \cdots & \mathbf{0}_{L \times 9N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{L \times 12N} & \mathbf{0}_{L \times 9N} & \cdots & \mathbf{\Theta} \end{bmatrix} \in \mathbb{R}^{LM \times 9NM}.$$

Here, $\mathbf{\Theta} \equiv [s_1, \dots, s_L]^T \in \mathbb{R}^{L \times 9N}$, where $l = 1, \dots, L$ and

$$(23) \quad s_l = [\mathbf{0}_{1 \times 3N}, -\mathbf{R}_l, -\mathbf{R}_l, -\mathbf{R}_l, \mathbf{0}_{1 \times 3N}]$$

is a $1 \times 9N$ row vector. The elements of the row vector $\mathbf{R}_l \in \mathbb{R}^{1 \times N}$ are assigned values to compute the pressure wavefield at the l^{th} transducer using trilinear interpolation.

By use of Eqns. (17), (19), and (22), the PACT imaging model in Eq. (11) can be expressed as

$$(24) \quad \hat{\mathbf{p}} = \mathbf{M} \mathbf{T}_{M-1} \cdots \mathbf{T}_1 \mathbf{T}_0 \mathbf{p}_0.$$

and therefore the system matrix is identified as

$$(25) \quad \mathbb{H} = \mathbf{M} \mathbf{T}_{M-1} \cdots \mathbf{T}_1 \mathbf{T}_0.$$

The explicit form of \mathbb{H}^\dagger is therefore given by

$$(26) \quad \mathbb{H}^\dagger = \mathbf{T}_0^\dagger \mathbf{T}_1^\dagger \cdots \mathbf{T}_{M-1}^\dagger \mathbf{M}^\dagger,$$

where the superscript \dagger denotes the conjugate transpose of a matrix.

4. Implementation of the forward and adjoint operators. Since a typical computational grid for simulating the stress tensor field that covers an entire human skull can consist of 500 million cells or more, there is a need for computationally efficient implementations of the forward and adjoint operators. To address this, a massively parallel implementation of the FDTD method based on NVIDIA's CUDA framework for general-purpose GPU computation was implemented [35, 36]. In this way, the action of both the operators \mathbb{H} and \mathbb{H}^\dagger could be computed by use of multiple GPUs utilizing the message passing interface (MPI) [36].

To prevent acoustic waves from reflecting off the edge of the simulation grid, an anisotropic absorbing boundary condition called a perfectly matched layer (PML) was implemented [4, 6, 5, 7]. For solving the photoacoustic wave equation in elastic, linear isotropic media, a special form of the PML called a convolutional-PML (C-PML) was implemented [26, 43]. To incorporate the C-PML, auxiliary memory variables need to be introduced [26, 43]. Due to the incorporation of the auxiliary memory variables, both \mathbb{H} and \mathbb{H}^\dagger need to be modified. The modified \mathbb{H} and \mathbb{H}^\dagger operators after the incorporation of the C-PML are described in the Appendix.

The action of the adjoint matrix was implemented according to Eqn. (26). It can be verified that $\mathbf{p}^{adj} = \mathbb{H}^\dagger \hat{\mathbf{p}}$ can be computed as

$$\begin{aligned} (27a) \quad \mathbf{v}_{M-1} &= \boldsymbol{\Theta}^T \hat{\mathbf{p}}_{M-1} \\ (27b) \quad \mathbf{v}_{m-1} &= \boldsymbol{\Theta}^T \hat{\mathbf{p}}_{m-1} + \mathbf{W}^T \mathbf{v}_m \\ (27c) \quad \mathbf{p}^{adj} &= \boldsymbol{\tau}^T \mathbf{v}_0. \end{aligned}$$

Moreover, the recursive temporal backward update step of Eqn. (27b) can be written in terms of the update of the field variables similar to Eqn. (10) as

$$\begin{aligned} (28a) \quad \tilde{\mathbf{u}}_{m-\frac{1}{2}} &= \bar{\mathbf{u}}_{m+\frac{1}{2}} + \boldsymbol{\Psi}^T \bar{\boldsymbol{\sigma}}_{m+1} \\ (28b) \quad \bar{\boldsymbol{\sigma}}_m &= \bar{\boldsymbol{\sigma}}_{m+1} + \boldsymbol{\Phi}^T \tilde{\mathbf{u}}_{m-\frac{1}{2}} + \mathcal{I}_2 \boldsymbol{\Theta}^T \hat{\mathbf{p}}_{m+1} \\ (28c) \quad \bar{\mathbf{u}}_{m-\frac{1}{2}} &= \mathbf{J} \tilde{\mathbf{u}}_{m-\frac{1}{2}} \end{aligned}$$

where $\mathcal{I}_2 \mathbf{v}_m \equiv \bar{\boldsymbol{\sigma}}_m$.

5. Validation studies. The implementation of the forward operator \mathbb{H} , as described by Eqn. (25), was validated by comparing the results obtained from the FDTD simulation with a known analytical solution. As shown in Figure 2a, the analytic solution was computed for a lossless semi-infinite medium, with fluid and linear isotropic solid media divided by a planar boundary. A monopole line source was placed in the fluid medium with the z -axis chosen normal to the interface between the solid/fluid media. Furthermore, the y -axis was chosen parallel to the line source located at $x = 0, z = h_T$. The receiving transducer, in this configuration, was located at $x = d, z = h_R$. In the setup described, the strength of the line source and the properties of the configuration were both independent of y . The analytical solution was computed via the Cagniard-De Hoop method [17, 16, 18].

In the FDTD simulation, a computational volume of $58.2 \text{ mm} \times 346.8 \text{ mm} \times 58.2 \text{ mm}$ was employed. Furthermore, the parameters of the configuration shown in Figure 2a were set as $h_T = 2.1 \text{ mm}$, $h_R = 2.1 \text{ mm}$ and $d = 2.1 \text{ mm}$, $\rho_s = 1200 \frac{\text{kg}}{\text{m}^3}$, $\lambda_s = 9.48 \text{ GPa}$, $\alpha_s = 0.0 \frac{1}{\mu\text{s}}$, $\mu_s = 2.352 \text{ GPa}$, $\rho_f = 1000 \frac{\text{kg}}{\text{m}^3}$, $\lambda_f = 2.25 \text{ GPa}$, $\alpha_f = 0.0 \frac{1}{\mu\text{s}}$, and $\mu_f = 0.0 \text{ GPa}$. An linear isotropic grid size of $\Delta x = 0.15 \text{ mm}$ was employed. The thickness of the C-PML was 4.5 mm on all sides of the 3-D computational grid. The pressure values obtained at the receiver locations were sampled with a sampling rate of 40 MHz . The digital representation of the line source had a Gaussian spread in the x - z -plane with a standard deviation of 1 mm . The results of the FDTD simulation are superimposed on the analytical solution in Figure 2b. The pressure profiles produced by the two methods are found to be nearly overlapping, indicating that the FDTD method possesses a high degree of accuracy.

In addition to validating the forward operator \mathbb{H} , the discrete adjoint operator was validated by application of the inner product test. The inner product test involves verifying the identity $\langle \mathbb{H}\mathbf{f}, \mathbf{g} \rangle_{\mathbb{V}} = \langle \mathbf{f}, \mathbb{H}^\dagger \mathbf{g} \rangle_{\mathbb{U}}$, where $\mathbf{f} \in \mathbb{U}$ and $\mathbf{g} \in \mathbb{V}$. Here, \mathbb{U} and \mathbb{V} represent the Euclidean spaces $\mathbb{R}^{N \times 1}$ and $\mathbb{R}^{LM \times 1}$, respectively. It was observed that the inner product test agreed to a 6-digit accuracy, thus validating the implementation of the discrete adjoint operator \mathbb{H}^\dagger .

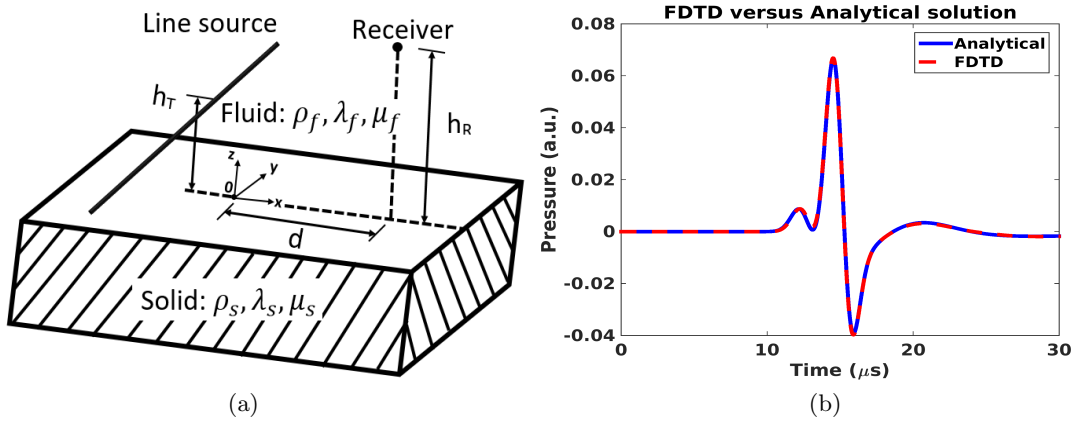


Figure 2. (a) The setup used to conduct the validation study. (b) A plot comparing the pressure profile obtained at the receiver location using the analytical solution and the FDTD simulation.

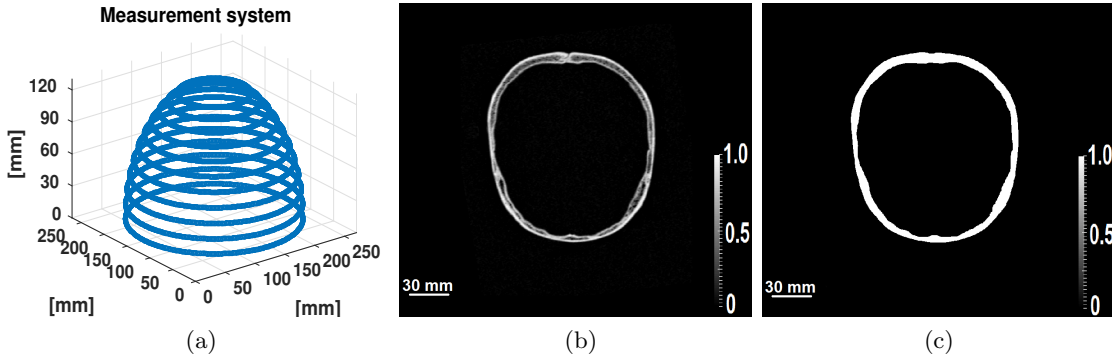


Figure 3. (a) The 3D scanning geometry used for the computer-simulation studies. (b) A 2D slice of X-ray CT image of the skull and (c) the corresponding mask generated by the segmentation algorithm.

6. Computer simulation studies. Computer-simulation studies were conducted in which the adjoint operator \mathbb{H}^\dagger was employed as a heuristic reconstruction operator. The performance of the adjoint operator was compared with a canonical backprojection (BP) reconstruction algorithm [28, 15] that assumed a homogeneous lossless fluid medium. To further study the impact of modeling shear wave propagation in 3D transcranial PACT, additional computer-simulation studies were conducted to assess the performance of the adjoint operator for cases where the shear modulus of the skull was assumed to be zero.

6.1. Methods.

6.1.1. Imaging geometry and phantom description. A 3D computational volume of $270.0 \text{ mm} \times 270.0 \text{ mm} \times 135.6 \text{ mm}$ was employed. The 3D scanning geometry, as shown in Figure 3a, consisted of 11 rings of varying radii with 400 transducers evenly distributed in each ring. The linear isotropic, elastic medium used in the simulation studies was generated from 3D X-ray CT images of a human skull. The intact human skull was purchased from

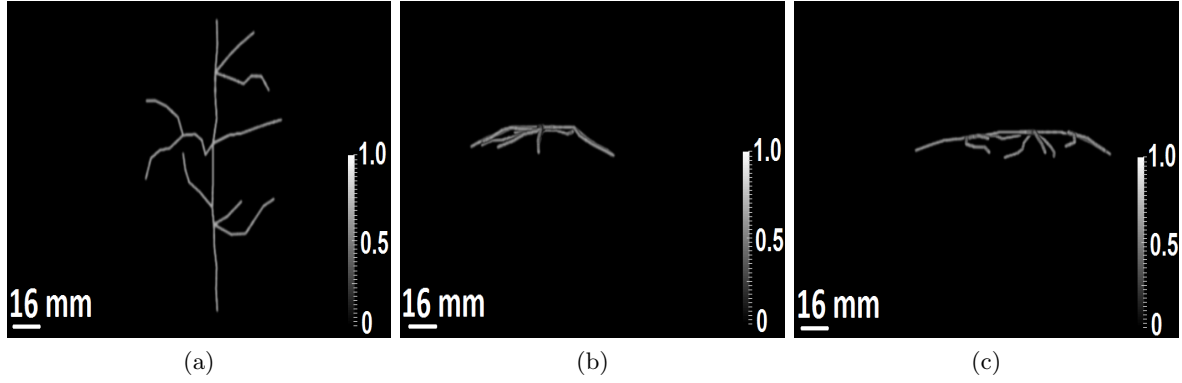


Figure 4. The maximum intensity projection of the initial pressure distribution (a) along the z -axis, (b) along the y -axis, and (c) along the x -axis.

Skull Unlimited International Inc. (Oklahoma City, OK) and was donated by an 83-year-old Caucasian male. The CT images were employed to infer the thickness and contour of the skull.

For the simulation studies involving \mathbb{H}^\dagger , we assumed the skull to be an acoustically homogeneous elastic linear isotropic medium. While we consider a relatively simple skull model, the proposed approach could also be applied for more complex skull models, such as those that consider the heterogeneity within the skull. In that case, more effort may be required to accurately estimate the acoustic properties of the skull. In order to extract the contour and location of the skull from CT images, a segmentation algorithm was employed. The segmentation algorithm generated a binary mask specifying the location of the skull within the 3D volume. A 2D slice of the CT image acquired from the human skull and the corresponding mask generated by use of the segmentation algorithm are shown in [Figure 3b](#) and [Figure 3c](#), respectively. The medium parameters in the 3D grid were assigned such that the skull acoustic parameters ($\rho = 1850 \frac{\text{kg}}{\text{m}^3}$, $c_l = 3.0 \frac{\text{mm}}{\mu\text{s}}$, $c_s = 1.5 \frac{\text{mm}}{\mu\text{s}}$ and $\alpha = 0.1 \frac{1}{\mu\text{s}}$) were set at all grid positions where mask was equal to one and the background acoustic parameters ($\rho = 1000 \frac{\text{kg}}{\text{m}^3}$, $c_l = 1.5 \frac{\text{mm}}{\mu\text{s}}$, $c_s = 0.0 \frac{\text{mm}}{\mu\text{s}}$ and $\alpha = 0.0 \frac{1}{\mu\text{s}}$) were set at all grid positions where mask was equal to zero. At the material interface between the skull and the background fluid medium, the density and the absorption values were arithmetically averaged to avoid any instability issues with the FDTD wave equation solver.

The initial pressure distribution assumed in the simulation studies mimicked cortical blood vessels (CBVs). The phantom, shown in [Figure 4](#), consisted of CBVs positioned approximately 6 mm below the inner surface of the skull. The 2D maximum intensity projection images along the x -, y - and z -axis of the initial pressure distribution are shown in [Figure 4](#).

6.1.2. Image reconstruction studies. To demonstrate the use of \mathbb{H}^\dagger as a reconstruction operator in 3D transcranial PACT, we conducted non-inverse crime computer-simulation studies [12], where different discretization strategies were employed to generate the measured data and to compute the action of \mathbb{H}^\dagger . In these studies, the forward data were generated using the phantom in [Figure 4](#) with a uniform grid size of $\Delta x = 0.225 \text{ mm}$ and a temporal sampling

rate of 50 MHz. Uncorrelated Gaussian noise was added to the simulated pressure signals. The standard deviation of the noise was set at 5% of the maximum signal value recorded. The action of \mathbb{H}^\dagger on the generated forward data was computed using a larger grid size of $\Delta x = 0.3$ mm. The simulated pressure data was not temporally downsampled before the application of the adjoint operator.

Images were also reconstructed from the noisy simulated data by use of the BP reconstruction algorithm. To find the optimal longitudinal SOS value for use in the BP reconstruction algorithm, we tuned the SOS over a range of values and picked the value that gave us the smallest mean squared error (MSE). The BP images were reconstructed using a uniform longitudinal SOS set to $1.540 \frac{mm}{\mu s}$, and using an uniform spatial grid of pitch $\Delta x = 0.3$ mm.

Finally, the importance of modeling shear wave propagation in the skull was further assessed by reconstructing images by use of a modified version of \mathbb{H}^\dagger , denoted as $\mathbb{H}_{\mu=0}^\dagger$, in which the shear modulus of the elastic medium was set to zero [22]. This corresponds to the un-physical situation in which the skull does not support shear wave propagation. In this simulation, a uniform grid size of $\Delta x = 0.3$ mm was used to compute the action of the adjoint operator. The longitudinal SOS of the skull was tuned over a range of values and the optimal value was selected based on the MSE. Note that, in this case, only the longitudinal SOS of the skull was tuned and the constant longitudinal SOS of the background fluid medium was fixed at $1.50 \frac{mm}{\mu s}$.

6.2. Computer-simulation: Results. The reconstructed images produced by use of \mathbb{H}^\dagger and the BP algorithm are shown in Figures 5a to 5c and Figures 5d to 5f. In both cases, the results were displayed as maximum intensity projection (MIP) images along three mutually perpendicular directions.

These results demonstrate that \mathbb{H}^\dagger can more effectively mitigate skull-induced image distortions than can the BP algorithm. Namely, despite being an approximate reconstruction operator, the image produced by application of \mathbb{H}^\dagger accurately displays the blood vessel geometry and possesses a much cleaner background and contains far fewer artifacts than the image reconstructed using the BP algorithm. It should also be noted that in cases for which \mathbb{H}^\dagger does not produce images of adequate quality, a more principled iterative approach to image reconstruction can be implemented by use of the operators \mathbb{H} and \mathbb{H}^\dagger .

The image reconstructed by use of $\mathbb{H}_{\mu=0}^\dagger$ is shown in Figures 5g to 5i. This image contains dramatically elevated artifact levels as compared to the one reconstructed by use of \mathbb{H}^\dagger , shown in Figures 5a to 5c. This demonstrates the importance of compensating for both the acoustic and the elastic properties of the skull in the reconstruction algorithm.

7. Experimental studies. Studies that utilized experimental PACT data produced by a physical phantom were also conducted.

7.1. Methods.

7.1.1. Imaging geometry and phantom description. A single element transducer was scanned over 4400 locations (11 rings with 400 evenly distributed positions per ring) in the configuration shown in Figure 3a to acquire the experimental data. A short 10 ns laser pulse (Nd-YAG Quantel Brilliant B laser with second harmonic generator) with a repetition rate of 10 Hz at a wavelength of 532 nm was used to irradiate a sample located in the center of

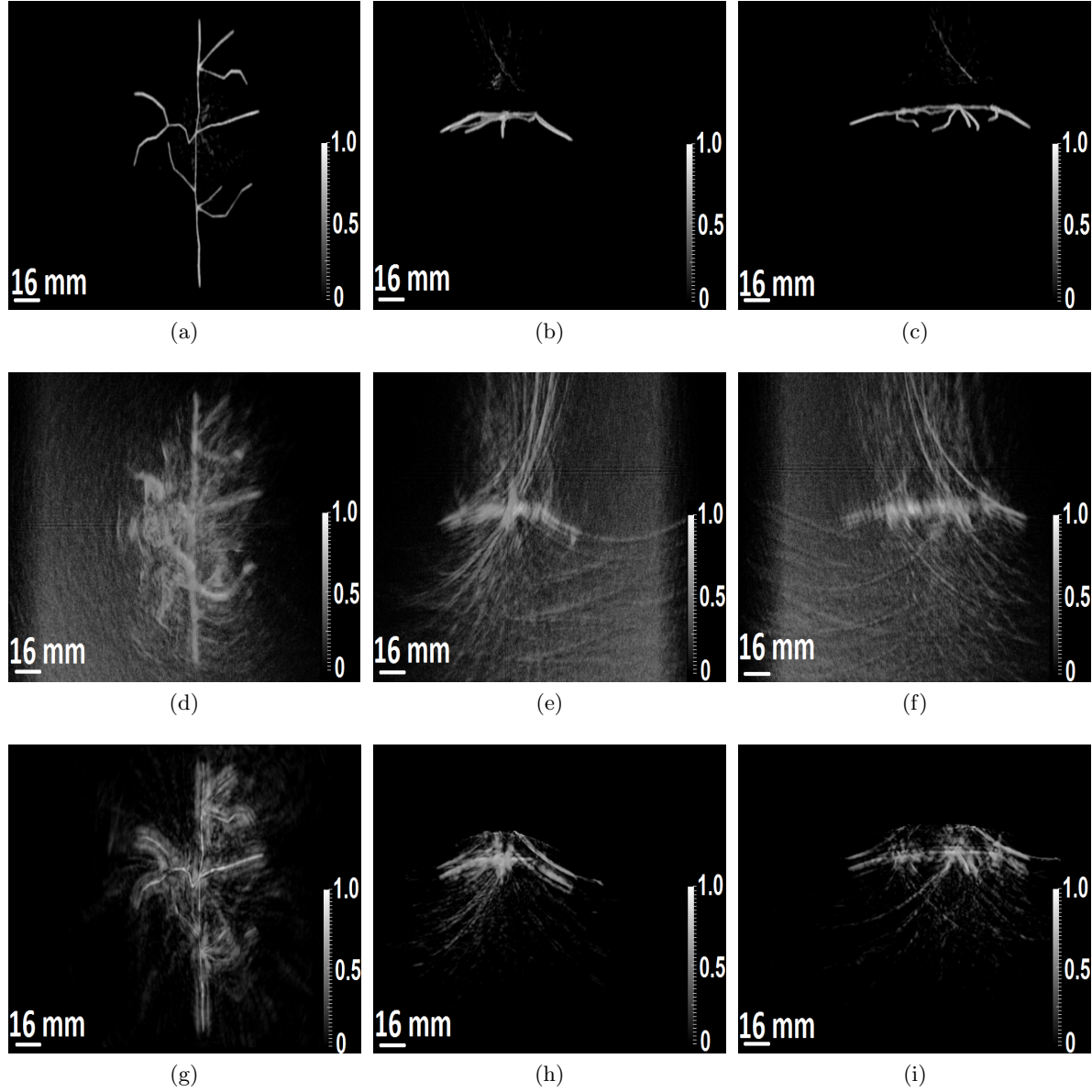


Figure 5. The maximum intensity projection of the reconstructed initial pressure distribution using \mathbb{H}^\dagger , along (a) the z-axis, (b) the y-axis, and (c) the x-axis. The maximum intensity projection of the reconstructed initial pressure distribution using BP algorithm along (d) the z-axis, (e) the y-axis, and (f) the x-axis. The maximum intensity projection of the reconstructed initial pressure distribution using $\mathbb{H}_{\mu=0}^\dagger$, along (g) the z-axis, (h) the y-axis, and (i) the x-axis.

the measurement system as shown in [Figure 6a](#). The subsequently generated acoustic signals were detected by unfocused transducers, with a center frequency of 1 MHz and a bandwidth of 80 %. The electrical signals recorded by the transducers were sampled at a temporal sampling

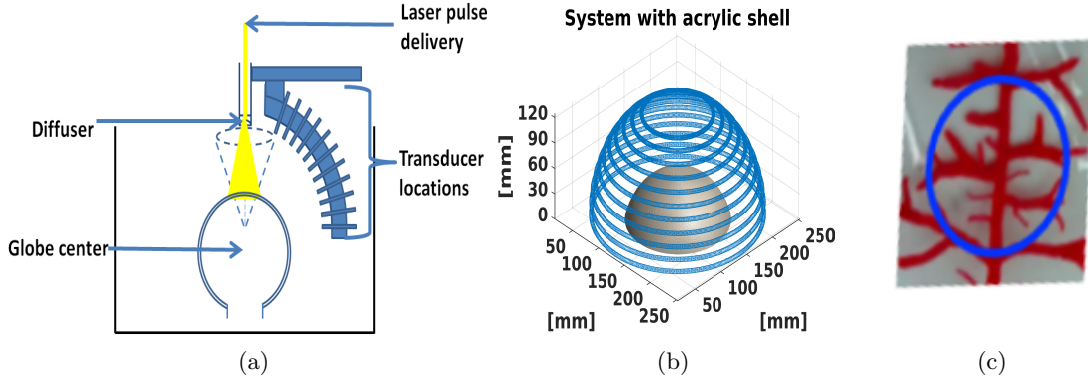


Figure 6. (a) The schematic of the system. (b) The position of the acrylic globe relative to the measurement system. (c) The vessels drawn on the inner surface of the acrylic globe.

rate of 20 MHz.

The system described above was employed to image a physical phantom whose design was motivated by transcranial PACT. The phantom was comprised of a spherical acrylic globe of thickness of 2.5 mm and an inner radius of 76.2 mm, placed within a 3D volume filled with water. The acrylic globe is an elastic solid that possesses SOS values that are representative of a human skull. The location of acrylic globe relative to the transducer array is shown in Figure 6b. Optically absorbing vessel-like structures were painted with Latex paint on the inner surface of the acrylic globe, as shown in Figure 6c. These vessels were intended to mimic cortical vessels that reside near the top surface of a brain.

In order to construct \mathbb{H}^\dagger , the acoustic parameters of the phantom need to be specified on a 3D Cartesian grid. The uniformly thick spherical acrylic shell (inner radius = 76.2 mm and thickness = 2.5 mm) was placed within the 3D volume such that the z-offset between the center of the shell and the first ring of transducer measurements was 30.2 mm. Thus, when computing \mathbb{H}^\dagger , only a spherical dome of the acrylic shell was included in the 3D simulation volume as shown by Figure 6b. The acoustic parameters of the globe were set to be $\rho = 1200 \frac{\text{kg}}{\text{m}^3}$, $c_l = 2.8 \frac{\text{m}}{\mu\text{s}}$, $c_s = 1.4 \frac{\text{mm}}{\mu\text{s}}$, and $\alpha = 0.1 \frac{1}{\mu\text{s}}$. In addition, the acoustic parameters of the homogeneous background (water bath) were specified as $\rho = 1000 \frac{\text{kg}}{\text{m}^3}$, $c_l = 1.5 \frac{\text{mm}}{\mu\text{s}}$, $c_s = 0.0 \frac{\text{mm}}{\mu\text{s}}$ and $\alpha = 0.0 \frac{1}{\mu\text{s}}$. Similar to the computer-simulation study, at the material interface between the globe and the background fluid medium, the density and the absorption values were arithmetically averaged to avoid any instability issues with the FDTD wave equation solver.

7.1.2. Data preprocessing. Prior to image reconstruction, the measured data were preprocessed. The preprocessing involved deconvolving the acquired data with the electrical impulse response (EIR) of the transducer. The measured EIR of the transducer is shown in Figure 7. The Wiener deconvolution method was employed to extract the deconvolved PA signals from the raw electrical transducer measurements. After deconvolution, the data were filtered with a Hann-window low-pass filter with a cutoff frequency of 2 MHz. The filtered data were also upsampled by a factor of 2.5, with the goal of circumventing numerical stability issues with the wave equation solver.

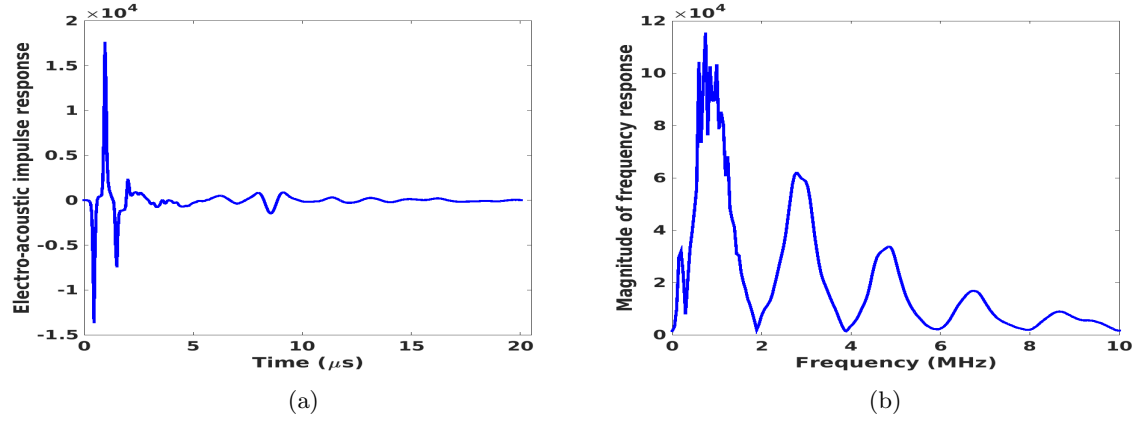


Figure 7. (a) The temporal profile of the measured EIR of the transducer. (b) The magnitude of the frequency response of the EIR of the transducer.

7.1.3. Image reconstruction studies. A 3D computational volume of $270 \text{ mm} \times 270 \text{ mm} \times 135.6 \text{ mm}$ was employed in the experimental studies. The action of \mathbb{H}^\dagger on the preprocessed data was computed with a uniform grid size of $\Delta x = 0.3 \text{ mm}$. The BP reconstruction algorithm was also applied to the preprocessed experimental data. The optimal longitudinal SOS value for the BP reconstruction algorithm was chosen by considering a range of values and selecting the value that gave us the best reconstructed image quality, as subjectively judged via visual inspection. For this study, the optimal longitudinal SOS of the acrylic globe was set to $1.650 \frac{\text{mm}}{\mu\text{s}}$. A uniform spatial grid size of $\Delta x = 0.3 \text{ mm}$ was used to compute the image reconstructed using the BP algorithm.

As in the computer-simulation studies, image were also reconstructed by use of the operator $\mathbb{H}_{\mu=0}^\dagger$. In this simulation, a uniform grid size of $\Delta x = 0.3 \text{ mm}$ was employed. Moreover, the longitudinal speed of sound of the acrylic globe was tuned over a range of values. Similar to the BP algorithm, the optimal longitudinal SOS of the acrylic globe was selected based on the reconstructed image quality.

7.2. Experimental studies: Results. The images reconstructed from the experimental data are shown in Figure 8. Figures 8a to 8c displays the image reconstructed by application of \mathbb{H}^\dagger , while the image reconstructed by use of the BP algorithm is shown in Figures 8d to 8f. In the BP reconstruction algorithm, the longitudinal speed of sound of the background fluid media was set to be $1.507 \frac{\text{mm}}{\mu\text{s}}$. These results demonstrate that the \mathbb{H}^\dagger can more effectively mitigate image distortions due to acrylic globe (i.e., simulated skull structure) than can the BP algorithm. Additionally, some of the smaller vessel structures are not identifiable in the BP image but are present in the image reconstructed by use of \mathbb{H}^\dagger .

The image reconstructed by use of $\mathbb{H}_{\mu=0}^\dagger$ is shown in Figures 8g to 8i. This image contains dramatically elevated artifact levels as compared to the one reconstructed by use of \mathbb{H}^\dagger , shown in Figures 8a to 8c. This again demonstrates that neglecting to model the elastic properties of the medium can lead to significant deterioration in image quality.

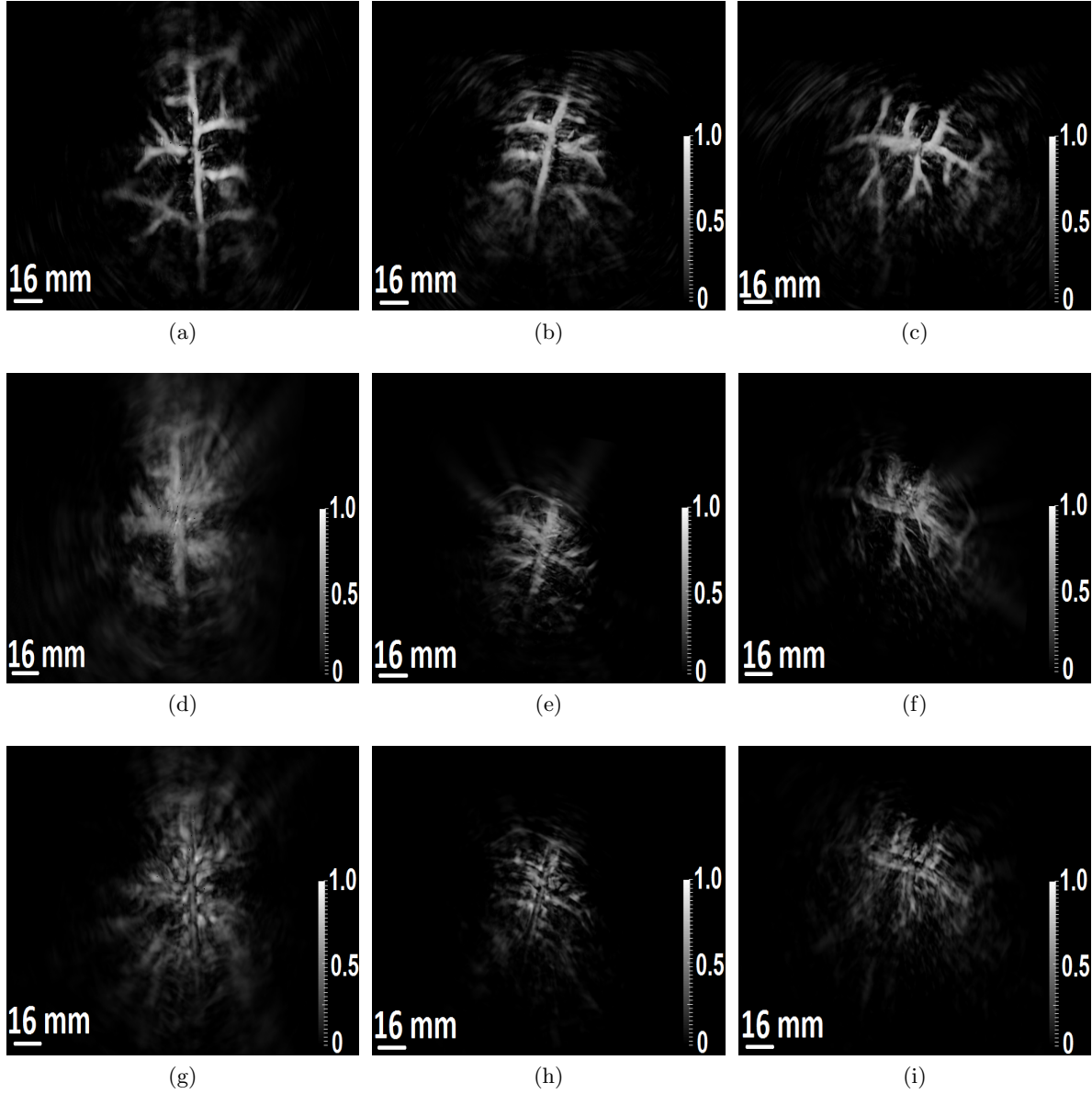


Figure 8. The reconstructed initial pressure distribution using \mathbb{H}^\dagger along three different views are shown in (a)-(c). The reconstructed initial pressure distribution using BP algorithm along three different views are shown in (d)-(f). The reconstructed initial pressure distribution using $\mathbb{H}^\dagger_{\mu=0}$ along three different views are shown in (g)-(i).

8. Conclusion. In this work, a forward and adjoint operator pair was introduced for use in transcranial PACT. To account for longitudinal-to-shear-wave conversions within the skull, these operators were based on a 3D elastic wave equation. Massively parallel implementations of these operators employing multiple graphics processing units (GPUs) were also developed.

The developed numerical framework was validated and investigated in computer-simulation and experimental phantom studies whose designs were motivated by transcranial PACT applications.

The numerical studies presented employed the adjoint operator \mathbb{H}^\dagger and a canonical BP reconstruction operator for image reconstruction. The two reconstruction approaches differ in that the operator \mathbb{H}^\dagger is based on the elastic wave equation for a specified heterogeneous elastic medium, while the canonical BP method assumes a homogeneous fluid medium. Our results demonstrated that neglecting to model the heterogeneous elastic properties of the medium can lead to significant deterioration in image quality. Although not presented, use of a filtered BP algorithm that was formed by discretizing an exact inversion formula [15, 52] did not alter the presented findings. In addition, the numerical and experimental results also demonstrated that the failure to model the shear wave propagation in elastic isotropic media can lead to significant deterioration in image quality.

There remain several important topics for further investigation. The proposed attenuation model is a diffusive model, where the attenuation coefficient is independent of the frequency. However, attenuation of elastic waves propagating in the skull is frequency-dependent. Therefore, the accuracy of proposed matched operator pair can be improved by incorporating a frequency dependent attenuation model. Furthermore, in both the computer-simulation and the experimental studies, to perform accurate reconstruction we require prior knowledge of the spatial distribution of the elastic parameters of the medium. Estimating the elastic parameters in an human skull is a challenging task, as the human skull is an acoustically heterogeneous medium. The skull is a three-layer structure with a porous zone, called the diploe, stacked between two dense layers, the outer and inner tables [2]. Recently, the possibility to deduce acoustic properties of the skull from adjunct CT data has gained some traction [33, 25, 2, 14]. As CT images provide information about the internal structure of the skull, it has been used to estimate the internal heterogeneities in density, speed and absorption. Hence, a study investigating the impact of errors in estimating acoustic parameters using adjunct CT data on the accuracy of the reconstructed image can be a subject of further study.

In addition, the use of \mathbb{H} and \mathbb{H}^\dagger in studies of iterative image reconstruction is a natural topic for investigation. The developed massively parallel implementations of \mathbb{H} and \mathbb{H}^\dagger using multiple GPUs will facilitate these studies. For example, the matched operator pair can be employed in an iterative image reconstruction method that seeks to minimize a penalized least square functional. Image reconstruction approaches such as this will be necessary when \mathbb{H}^\dagger does not produce useful images.

Appendix- Convolutional PML (C-PML).

8.1. Introduction to C-PML. One of the most popular methods of implementing absorbing boundary conditions is the PML. It was originally developed by Berenger for use with Maxwell's equations using a split-field formulation [4, 6]. One of the drawbacks to the split-field formulation of the classical PML is that it increases the number of field variables that need to be stored and the number of differential equations that need to be solved over the whole domain.

Later, a complex coordinate stretching technique was developed [43, 11], which is equivalent to Berenger's formulation in the frequency-domain, but which has a much more efficient

implementation in the time-domain. In this approach, the spatial derivatives along each direction are replaced with a scaled version. For example, in the x -direction, the scaled spatial derivative is given by

$$(29) \quad \partial_{\tilde{x}} = \frac{1}{S_x(x, \omega)} \partial_x,$$

where $S_x(x, \omega)$ is a spatially-dependent, complex scaling factor given in the temporal frequency domain by

$$(30) \quad S_x(x, \omega) = 1 + \frac{d_x(x)}{j\omega}.$$

Here, $d_x(x)$ is the profile for the absorption of the absorbing boundary layer, which is equal to zero outside of the layer. In the time-domain, Eqn. (29) is given by

$$(31) \quad \partial_{\tilde{x}} = s_x(x, t) *_t \partial_x,$$

which is why this method is also sometimes called a convolutional PML (C-PML).

While Berenger's formulation of the PML gives rise to an absorbing boundary layer with a reflection coefficient of zero for all angles of incidence, this property only holds in the continuous case. Upon discretization, for waves arriving at grazing incidence, a large amount of energy is sent back into the main domain in the form of spurious reflected waves. This makes the discrete classical PML less efficient for cases where the sources are located close to the edge of the grid, which is commonly encountered in elastic wave modeling [26]. To circumvent this problem, a more general form for the scaling factor was developed [29, 43], given by

$$(32) \quad S_x(x, \omega) = \kappa_x(x) + \frac{d_x(x)}{\beta_x(x) + j\omega}.$$

Given Eqns. (29), (31), and (32), one can write

$$(33) \quad s_x(x, t) = \frac{1}{\kappa_x(x)} \delta(t) + \zeta_x(x, t),$$

where

$$(34) \quad \zeta_x(x, t) \equiv \frac{d_x(x)}{\kappa_x(x)^2} \exp(-[d_x(x)/\kappa_x(x) + \beta_x(x)]t) U(t),$$

and $U(t)$ denotes the Heaviside step function. Thus, our modified spatial derivative operator is given by

$$(35) \quad \partial_{\tilde{x}} = \frac{1}{\kappa_x(x)} \partial_x + \zeta_x(t) *_t \partial_x.$$

From a numerical point of view, the calculation of the convolution is costly because it requires at each time step to sum over all previous time steps. However, given the particular form of

ζ_i , it is possible to compute the convolution using a recursive technique by use of a memory variable [31]. Given a memory variable ϕ_x , the derivative of a field variable along the x^{th} direction can be written as

$$(36) \quad \partial_{\tilde{x}} = \frac{1}{\kappa_x} \partial_x + \phi_x,$$

where this memory variable is updated recursively in time according to

$$(37) \quad \phi_x^n = b_x \phi_x^{n-1} + a_x (\partial_x)^{n-\frac{1}{2}},$$

where the spatial derivative acts on the specific field variable associated with the memory variable $\phi_x(x)$. In addition, the PML decay constants $a_x(x)$ and $b_x(x)$ in Eqn. (37) are given by

$$(38a) \quad b_x(x) = \exp(-(d_x(x)/\kappa_x(x) + \beta_x(x)) \Delta t)$$

$$(38b) \quad a_x(x) = \frac{d_x}{\kappa_x(d_x + \kappa_x \beta_x(x))} (b_x(x) - 1).$$

8.2. Application of C-PML to elastic wave equation. In order to apply the C-PML-based absorbing boundary conditions to the elastic wave equation in Eqn. (3), one can replace the spatial derivative operations defined in Eqn. (3), with the modified spatial derivative operations defined in Eqn. (36). For simplicity, we consider $\kappa = 1$. The parameter κ was introduced to attenuate evanescent waves when modeling Maxwell's equations and may be less critical to elastic wave simulations [26]. Thus, given the modified spatial derivative operators in Eqn. (36), the collection of differential equations defined in Eqn. (3) can be written as

$$(39a) \quad \partial_t \dot{u}^1 + \alpha \dot{u}^1 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{1i} + \zeta_i * \partial_i \sigma^{1i}),$$

$$(39b) \quad \partial_t \dot{u}^2 + \alpha \dot{u}^2 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{2i} + \zeta_i * \partial_i \sigma^{2i}),$$

$$(39c) \quad \partial_t \dot{u}^3 + \alpha \dot{u}^3 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{3i} + \zeta_i * \partial_i \sigma^{3i}),$$

$$(39d) \quad \partial_t \sigma^{11} = \lambda \sum_{i=1}^3 (\partial_i u^i + \zeta_i * \partial_i u^i) + 2\mu (\partial_1 \dot{u}^1 + \zeta_1 * \partial_1 \dot{u}^1),$$

$$(39e) \quad \partial_t \sigma^{22} = \lambda \sum_{i=1}^3 (\partial_i u^i + \zeta_i * \partial_i u^i) + 2\mu (\partial_2 \dot{u}^2 + \zeta_2 * \partial_2 \dot{u}^2),$$

$$(39f) \quad \partial_t \sigma^{33} = \lambda \sum_{i=1}^3 (\partial_i u^i + \zeta_i * \partial_i u^i) + 2\mu (\partial_3 \dot{u}^3 + \zeta_3 * \partial_3 \dot{u}^3),$$

$$(39g) \quad \partial_t \sigma^{23} = \mu (\partial_2 \dot{u}^3 + \partial_3 \dot{u}^2 + \zeta_2 * \partial_2 \dot{u}^3 + \zeta_3 * \partial_3 \dot{u}^2),$$

$$(39h) \quad \partial_t \sigma^{13} = \mu (\partial_1 \dot{u}^3 + \partial_3 \dot{u}^1 + \zeta_1 * \partial_1 \dot{u}^3 + \zeta_3 * \partial_3 \dot{u}^1) \text{ and}$$

$$(39i) \quad \partial_t \sigma^{12} = \mu (\partial_1 \dot{u}^2 + \partial_2 \dot{u}^1 + \zeta_1 * \partial_1 \dot{u}^2 + \zeta_2 * \partial_2 \dot{u}^1).$$

Note that, because the elastic media is linear isotropic ($\sigma^{ij} = \sigma^{ji}$), it is not necessary to compute all nine components of the stress tensor. Let the memory variables be defined as

$$(40a) \quad \phi^{ij} \leftarrow \zeta_j * \partial_j \sigma^{ij}$$

$$(40b) \quad \psi^{ij} \leftarrow \zeta_j * \partial_j u^i,$$

where the memory variables are updated according to

$$(41a) \quad \phi_{m+\frac{1}{2}}^{ij} = b_j (r^{ij}) \phi_{m-\frac{1}{2}}^{ij} + a_j (r^{ij}) \partial_j \sigma_m^{ij}$$

$$(41b) \quad \psi_m^{ij} = b_j (r^i) \psi_{m-1}^{ij} + a_j (r^i) \partial_j u_{m-\frac{1}{2}}^i.$$

Substituting the memory variables defined in Eqn. (40) into Eqn. (39) we have

$$(42a) \quad \partial_t \dot{u}^1 + \alpha \dot{u}^1 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{1i} + \phi^{1i}),$$

$$(42b) \quad \partial_t \dot{u}^2 + \alpha \dot{u}^2 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{2i} + \phi^{2i}),$$

$$(42c) \quad \partial_t \dot{u}^3 + \alpha \dot{u}^3 = \frac{1}{\rho} \sum_{i=1}^3 (\partial_i \sigma^{3i} + \phi^{3i}),$$

$$(42d) \quad \partial_t \sigma^{11} = \lambda \sum_{i=1}^3 (\partial_i u^i + \psi^{ii}) + 2\mu (\partial_1 \dot{u}^1 + \psi^{11}),$$

$$(42e) \quad \partial_t \sigma^{22} = \lambda \sum_{i=1}^3 (\partial_i u^i + \psi^{ii}) + 2\mu (\partial_2 \dot{u}^2 + \psi^{22}),$$

$$(42f) \quad \partial_t \sigma^{33} = \lambda \sum_{i=1}^3 (\partial_i u^i + \psi^{ii}) + 2\mu (\partial_3 \dot{u}^3 + \psi^{33}),$$

$$(42g) \quad \partial_t \sigma^{23} = \mu (\partial_2 \dot{u}^3 + \partial_3 \dot{u}^2 + \psi^{23} + \psi^{32}),$$

$$(42h) \quad \partial_t \sigma^{13} = \mu (\partial_1 \dot{u}^3 + \partial_3 \dot{u}^1 + \psi^{13} + \psi^{31}), \text{ and}$$

$$(42i) \quad \partial_t \sigma^{12} = \mu (\partial_1 \dot{u}^2 + \partial_2 \dot{u}^1 + \psi^{12} + \psi^{21}).$$

8.3. Discrete formulation. In order to explicitly write the forward and the adjoint operator for the FDTD elastic wave equation solver with the C-PML, we need to define the

following matrices:

$$(43a) \quad \mathcal{A}_{ij} \equiv \text{diag} \left(a_i \left(\mathbf{r}_1^j \right), \dots, a_i \left(\mathbf{r}_N^j \right) \right)$$

$$(43b) \quad \tilde{\mathcal{A}}_{ij} \equiv \text{diag} \left(a_i \left(\mathbf{r}_1^{ij} \right), \dots, a_i \left(\mathbf{r}_N^{ij} \right) \right)$$

$$(43c) \quad \mathcal{B}_{ij} \equiv \text{diag} \left(b_i \left(\mathbf{r}_1^j \right), \dots, b_i \left(\mathbf{r}_N^j \right) \right)$$

$$(43d) \quad \tilde{\mathcal{B}}_{ij} \equiv \text{diag} \left(b_i \left(\mathbf{r}_1^{ij} \right), \dots, b_i \left(\mathbf{r}_N^{ij} \right) \right)$$

$$(43e) \quad \mathcal{X}_{ij} \equiv \text{diag} \left(\chi^i \left(\mathbf{r}_1^j \right), \dots, \chi^i \left(\mathbf{r}_N^j \right) \right)$$

$$(43f) \quad \tilde{\mathcal{X}}_{ij} \equiv \text{diag} \left(\chi^i \left(\mathbf{r}_1^{ij} \right), \dots, \chi^i \left(\mathbf{r}_N^{ij} \right) \right),$$

where $a_i(\mathbf{r})$, $b_i(\mathbf{r})$ are PML decay constants in the i^{th} direction and $\chi^i(\mathbf{r})$ represents the PML indicator function which is one if the component of \mathbf{r} along the i^{th} direction is in the PML. Let the PML memory variable for the spatial derivative of σ^{ij} along the j -th direction be denoted as ϕ^{ij} , and the PML memory variable for the spatial derivative of $\dot{\mathbf{u}}^i$ along the j -th direction be denoted as ψ^{ij} . It should be noted that ϕ^{ij} is sampled at \mathbf{r}^j and ψ^{ij} at \mathbf{r}^{ij} . Let us define

$$(44a) \quad \bar{\phi}_m \equiv [\phi_m^{11}, \phi_m^{12}, \phi_m^{13}, \phi_m^{21}, \phi_m^{22}, \phi_m^{23}, \phi_m^{31}, \phi_m^{32}, \phi_m^{33}]^T$$

$$(44b) \quad \bar{\psi}_m \equiv [\psi_m^{11}, \psi_m^{12}, \psi_m^{13}, \psi_m^{21}, \psi_m^{22}, \psi_m^{23}, \psi_m^{31}, \psi_m^{32}, \psi_m^{33}]^T,$$

where $\bar{\phi}_m \in \mathbb{R}^{9N \times 1}$ and $\bar{\psi}_m \in \mathbb{R}^{9N \times 1}$.

Further, let us define the operators

$$(45a) \quad \mathbf{B}_{ij} \bar{\phi}_m \equiv \mathcal{B}_{ij} \phi_m^{ij},$$

$$(45b) \quad \tilde{\mathbf{B}}_{ij} \bar{\psi}_m \equiv \tilde{\mathcal{B}}_{ij} \psi_m^{ij},$$

$$(45c) \quad \mathbf{E}_{ij} \bar{\sigma}_m \equiv \mathcal{A}_{ij} \partial_j \sigma_m^{ij},$$

$$(45d) \quad \tilde{\mathbf{E}}_{ij} \bar{\mathbf{u}}_m \equiv \tilde{\mathcal{A}}_{ij} \partial_j \dot{\mathbf{u}}_m^i,$$

$$(45e) \quad \Gamma_i \bar{\phi}_m \equiv \Delta t \mathbf{Q}_i^{-1} \sum_{j=1}^3 \mathcal{X}_{ij} \phi_m^{ij}, \text{ and}$$

$$(45f) \quad \Delta_{ij} \bar{\psi}_m \equiv \Delta t \left[\delta_{ij} \Lambda_{ij} \sum_{k=1}^3 \psi_m^{kk} + \mathcal{M}_{ij} \left(\tilde{\mathcal{X}}_{ij} \psi_m^{ij} + \tilde{\mathcal{X}}_{ji} \psi_m^{ji} \right) \right],$$

where $\mathbf{B}_{ij}, \tilde{\mathbf{B}}_{ij}, \Gamma_i, \Delta_{ij} \in \mathbb{R}^{N \times 9N}$, $\mathbf{E}_{ij} \in \mathbb{R}^{N \times 6N}$, and $\tilde{\mathbf{E}}_{ij} \in \mathbb{R}^{N \times 3N}$. Given the definitions in Eqns. (43)–(45), we can write the discretized form of Eqn. (42) as

$$(46a) \quad \phi_{m+\frac{1}{2}}^{ij} = \mathbf{B}_{ij} \bar{\phi}_{m-\frac{1}{2}} + \mathbf{E}_{ij} \bar{\sigma}_m$$

$$(46b) \quad \dot{\mathbf{u}}_{m+\frac{1}{2}}^i = \mathbf{J}_i \bar{\mathbf{u}}_{m-\frac{1}{2}} + \Phi_i \bar{\sigma}_m + \Gamma_i \bar{\phi}_{m+\frac{1}{2}}$$

$$(46c) \quad \psi_{m+1}^{ij} = \tilde{\mathbf{B}}_{ij} \bar{\psi}_m + \tilde{\mathbf{E}}_{ij} \bar{\mathbf{u}}_{m+\frac{1}{2}}$$

$$(46d) \quad \sigma_{m+1}^{ij} = \sigma_m^{ij} + \Psi_{ij} \bar{\mathbf{u}}_{m+\frac{1}{2}} + \Delta_{ij} \bar{\psi}_{m+1}.$$

Further, let us also define

$$(47a) \quad \mathbf{B} \equiv [\mathbf{B}_{11}, \mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{21}, \mathbf{B}_{22}, \mathbf{B}_{23}, \mathbf{B}_{31}, \mathbf{B}_{32}, \mathbf{B}_{33}]^T,$$

$$(47b) \quad \tilde{\mathbf{B}} \equiv [\tilde{\mathbf{B}}_{11}, \tilde{\mathbf{B}}_{12}, \tilde{\mathbf{B}}_{13}, \tilde{\mathbf{B}}_{21}, \tilde{\mathbf{B}}_{22}, \tilde{\mathbf{B}}_{23}, \tilde{\mathbf{B}}_{31}, \tilde{\mathbf{B}}_{32}, \tilde{\mathbf{B}}_{33}]^T,$$

$$(47c) \quad \mathbf{E} \equiv [\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{13}, \mathbf{E}_{21}, \mathbf{E}_{22}, \mathbf{E}_{23}, \mathbf{E}_{31}, \mathbf{E}_{32}, \mathbf{E}_{33}]^T,$$

$$(47d) \quad \tilde{\mathbf{E}} \equiv [\tilde{\mathbf{E}}_{11}, \tilde{\mathbf{E}}_{12}, \tilde{\mathbf{E}}_{13}, \tilde{\mathbf{E}}_{21}, \tilde{\mathbf{E}}_{22}, \tilde{\mathbf{E}}_{23}, \tilde{\mathbf{E}}_{31}, \tilde{\mathbf{E}}_{32}, \tilde{\mathbf{E}}_{33}]^T,$$

$$(47e) \quad \mathbf{\Gamma} \equiv [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3]^T,$$

$$(47f) \quad \mathbf{\Delta} \equiv [\mathbf{\Delta}_{11}, \mathbf{\Delta}_{22}, \mathbf{\Delta}_{33}, \mathbf{\Delta}_{23}, \mathbf{\Delta}_{13}, \mathbf{\Delta}_{12}]^T,$$

$$(47g) \quad \mathbf{F} \equiv \mathbf{\Phi} + \mathbf{\Gamma}\mathbf{E}, \text{ and}$$

$$(47h) \quad \mathbf{G} \equiv \mathbf{\Psi} + \mathbf{\Delta}\tilde{\mathbf{E}},$$

where $\mathbf{B}, \tilde{\mathbf{B}} \in \mathbb{R}^{9N \times 9N}$, $\mathbf{E} \in \mathbb{R}^{9N \times 6N}$, $\tilde{\mathbf{E}} \in \mathbb{R}^{9N \times 3N}$, $\mathbf{\Gamma} \in \mathbb{R}^{3N \times 9N}$, $\mathbf{F} \in \mathbb{R}^{3N \times 6N}$, $\mathbf{G} \in \mathbb{R}^{6N \times 3N}$ and $\mathbf{\Delta} \in \mathbb{R}^{6N \times 9N}$. In the presence of the C-PML, the single matrix equation to determine the updated wavefield variables after time Δt is given by

$$(48) \quad \mathbf{v}'_{m+1} = \mathbf{W}' \mathbf{v}'_m,$$

where $\mathbf{v}'_m = \begin{bmatrix} \bar{\phi}_{m-\frac{1}{2}} \\ \bar{\mathbf{u}}_{m-\frac{1}{2}} \\ \bar{\psi}_m \\ \bar{\sigma}_m \end{bmatrix}$, and the propagator matrix $\mathbf{W}' \in \mathbb{R}^{27N \times 27N}$ is given by

$$(49) \quad \mathbf{W}' = \begin{bmatrix} \mathbf{B} & \mathbf{0}_{9N \times 3N} & \mathbf{0}_{9N \times 9N} & \mathbf{E} \\ \mathbf{\Gamma}\mathbf{B} & \mathbf{J} & \mathbf{0}_{3N \times 9N} & \mathbf{F} \\ \tilde{\mathbf{E}}\mathbf{\Gamma}\mathbf{B} & \tilde{\mathbf{E}}\mathbf{J} & \tilde{\mathbf{B}} & \tilde{\mathbf{E}}\mathbf{F} \\ \mathbf{G}\mathbf{\Gamma}\mathbf{B} & \mathbf{G}\mathbf{J} & \mathbf{\Delta}\tilde{\mathbf{B}} & \mathbf{I}_{6N \times 6N} + \mathbf{G}\mathbf{F} \end{bmatrix}.$$

Hence, the wavefield quantities can be propagated forward in time from $t = 0$ to $t = (M - 1)\Delta t$ as

$$(50) \quad \begin{bmatrix} \mathbf{v}'_0 \\ \mathbf{v}'_1 \\ \vdots \\ \mathbf{v}'_{M-1} \end{bmatrix} = \mathbf{T}'_{M-1} \cdots \mathbf{T}'_1 \begin{bmatrix} \mathbf{v}'_0 \\ \mathbf{0}_{27N \times 1} \\ \vdots \\ \mathbf{0}_{27N \times 1} \end{bmatrix},$$

where the $27NM \times 27NM$ matrices $\mathbf{T}'_m (m = 1, \dots, M - 1)$ are defined in terms of \mathbf{W}' as

$$(51) \quad \mathbf{T}'_m \equiv \begin{bmatrix} \mathbf{I}_{27N \times 27N} & \cdots & \mathbf{0}_{27N \times 27N} \\ \vdots & \ddots & \vdots & \mathbf{0}_{(m+1) \cdot 27N \times (M-m) \cdot 27N} \\ \mathbf{0}_{27N \times 27N} & \cdots & \mathbf{I}_{27N \times 27N} \\ \mathbf{0}_{27N \times 27N} & \cdots & \mathbf{W}' \\ \mathbf{0}_{(M-m-1) \cdot 27N \times m \cdot 27N} & & & \mathbf{0}_{(M-m-1) \cdot 27N \times (M-m) \cdot 27N} \end{bmatrix},$$

with \mathbf{W}' residing between the $(27N(m-1)+1)^{th}$ to $27Nm^{th}$ column and the $(27Nm+1)^{th}$ to $27N(m+1)^{th}$ rows of \mathbf{T}'_m . From the equation of state in Eqn. (3a), (3b) and the initial conditions in Eqn. (3c), the vector $(\mathbf{v}'_0, \mathbf{0}_{27N \times 1}, \dots, \mathbf{0}_{27N \times 1})^T$ can be computed from the initial pressure distribution \mathbf{p}_0 as

$$(52) \quad \begin{bmatrix} \mathbf{v}'_0 \\ \mathbf{0}_{27N \times 1} \\ \vdots \\ \mathbf{0}_{27N \times 1} \end{bmatrix} = \mathbf{T}'_0 \mathbf{p}_0,$$

where

$$(53) \quad \mathbf{T}'_0 \equiv [\boldsymbol{\tau}', \mathbf{0}_{27N \times N} \cdots, \mathbf{0}_{27N \times N}]^T \in \mathbb{R}^{27NM \times N} \text{ and}$$

$$(54) \quad \boldsymbol{\tau}' \equiv [\mathbf{0}_{21N \times N}, -\mathbf{I}_{N \times N}, -\mathbf{I}_{N \times N}, -\mathbf{I}_{N \times N}, \mathbf{0}_{3N \times N}]^T \in \mathbb{R}^{27N \times N}$$

and \mathbf{p}_0 is defined by Eqn. (13). Note that the memory variable associated with the spatial derivative of the stress tensor is set to zero initially. This implies that the support of the initial pressure distribution should be at least 2 grid positions away from the PML in all three directions.

In general, the transducer locations \mathbf{r}_l^d at which the data $\hat{\mathbf{p}}$ are recorded will not coincide with the vertices of the Cartesian grid at which the propagated field quantities are computed. The measured data $\hat{\mathbf{p}}$ can be related to the computed field quantities via an interpolation operation defined as

$$(55) \quad \hat{\mathbf{p}} = \mathbf{M}' \begin{bmatrix} \mathbf{v}'_0 \\ \mathbf{v}'_1 \\ \vdots \\ \mathbf{v}'_{M-1} \end{bmatrix}, \text{ where } \mathbf{M}' \equiv \begin{bmatrix} \boldsymbol{\Theta}' & \mathbf{0}_{L \times 27N} & \cdots & \mathbf{0}_{L \times 27N} \\ \mathbf{0}_{L \times 27N} & \boldsymbol{\Theta}' & \cdots & \mathbf{0}_{L \times 27N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{L \times 27N} & \mathbf{0}_{L \times 27N} & \cdots & \boldsymbol{\Theta}' \end{bmatrix} \in \mathbb{R}^{LM \times 27N}.$$

Here, $\boldsymbol{\Theta}' \equiv [s'_1, \dots, s'_L]^T \in \mathbb{R}^{L \times 27N}$, where $l = 1, \dots, L$ and

$$(56) \quad s'_l = [\mathbf{0}_{1 \times 21N}, -\mathbf{R}_l, -\mathbf{R}_l, -\mathbf{R}_l, \mathbf{0}_{1 \times 3N}]$$

is a $1 \times 27N$ row vector. The elements of row vector $\mathbf{R}_l \in \mathbb{R}^{1 \times N}$ are assigned values to compute the pressure at the l^{th} transducer using trilinear interpolation. Thus, from Eqn. (50), (52), (55), the explicit form of the system matrix that solves the initial value problem of Eqn. (3) in a discrete setting with a C-PML is given by

$$(57) \quad \hat{\mathbf{p}} = \mathbf{M}' \mathbf{T}'_{M-1} \cdots \mathbf{T}'_1 \mathbf{T}'_0 \mathbf{p}_0.$$

Comparing Eqn. (57) with Eqn. (11) the explicit form of the system matrix is given by

$$(58) \quad \mathbb{H} = \mathbf{M}' \mathbf{T}'_{M-1} \cdots \mathbf{T}'_1 \mathbf{T}'_0.$$

Hence, the explicit form of \mathbb{H}^\dagger is given by

$$(59) \quad \mathbb{H}^\dagger = \mathbf{T}'_0{}^\dagger \mathbf{T}'_1{}^\dagger \cdots \mathbf{T}'_{M-1}{}^\dagger \mathbf{M}'^\dagger.$$

The action of the adjoint matrix on the measured pressure data $\hat{\mathbf{p}}$ was implemented according to Eqn. (59). The state equations for computing $\mathbf{p}^{adj} = \mathbb{H}^\dagger \hat{\mathbf{p}}$ with the incorporation of the C-PML can be written as,

$$\begin{aligned} (60a) \quad \mathbf{v}'_{M-1} &= \boldsymbol{\Theta}'^T \hat{\mathbf{p}}_{M-1} \\ (60b) \quad \mathbf{v}'_m &= \boldsymbol{\Theta}'^T \hat{\mathbf{p}}_{m-1} + \mathbf{W}'^T \mathbf{v}'_m \\ m &= M-1, \dots, 1 \\ (60c) \quad \mathbf{p}^{adj} &= \boldsymbol{\tau}'^T \mathbf{v}'_0. \end{aligned}$$

Similar to Eqn. (28), the recursive temporal backward update step of Eqn. (60b) can be written as

$$\begin{aligned} (61a) \quad \tilde{\psi}_{m-1} &= \bar{\psi}_{m+1} + \boldsymbol{\Delta}^\dagger \bar{\sigma}_{m+1} \\ (61b) \quad \tilde{\mathbf{u}}_{m-\frac{1}{2}} &= \bar{\mathbf{u}}_{m+\frac{1}{2}} + \tilde{\mathbf{E}}^\dagger \bar{\psi}_{m+1} + \mathbf{G}^\dagger \bar{\sigma}_{m+1} \\ (61c) \quad \tilde{\phi}_{m-\frac{1}{2}} &= \bar{\phi}_{m+\frac{1}{2}} + \boldsymbol{\Gamma}^\dagger \tilde{\mathbf{u}}_{m-\frac{1}{2}} \\ (61d) \quad \bar{\sigma}_m &= \bar{\sigma}_{m+1} + \mathbf{E}^\dagger \bar{\phi}_{m+\frac{1}{2}} + \mathbf{F}^\dagger \tilde{\mathbf{u}}_{m-\frac{1}{2}} + \mathcal{I}'_4 \boldsymbol{\Theta}'^T \hat{\mathbf{p}}_{m+1} \\ (61e) \quad \bar{\mathbf{u}}_{m-\frac{1}{2}} &= \mathbf{J} \tilde{\mathbf{u}}_{m-\frac{1}{2}} \\ (61f) \quad \bar{\psi}_{m-1} &= \tilde{\mathbf{B}} \tilde{\psi}_{m-1} \\ (61g) \quad \bar{\phi}_{m-\frac{1}{2}} &= \mathbf{B} \tilde{\phi}_{m-\frac{1}{2}}, \end{aligned}$$

where $\mathcal{I}'_4 \mathbf{v}'_m \equiv \bar{\sigma}_m$.

Acknowledgments. This work was supported in part by NIH award EB01696301, 5T32EB01485505 and NSF award DMS1614305.

REFERENCES

- [1] Z. ALTERMAN AND F. K. JR., *Propagation of elastic waves in layered media by finite difference methods*, Bulletin of the Seismological Society of America, 58 (1968), pp. 367–398.
- [2] J.-F. AUBRY, M. TANTER, M. PERNOT, J.-L. THOMAS, AND M. FINK, *Experimental demonstration of noninvasive transskull adaptive focusing based on prior computed tomography scans*, The Journal of the Acoustical Society of America, 113 (2003), pp. 84–93.
- [3] A. BECK, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*, vol. 19, SIAM, 2014.
- [4] J.-P. BÉRENGER, *A perfectly matched layer for the absorption of electromagnetic waves*, Journal of Computational Physics, 114 (1994), pp. 185–200.
- [5] J.-P. BÉRENGER, *Perfectly matched layer for the fdtd solution of wave-structure interaction problems*, IEEE. Trans. Antennas Propogat., 44 (1996), pp. 110–117.
- [6] J.-P. BÉRENGER, *Three-dimensional perfectly matched layer for the absorption of electromagnetic waves*, Journal of Computational Physics, 127 (1996), pp. 363–379.
- [7] J.-P. BÉRENGER, *Improved pml for the fdtd solution of wave-structure interaction problems*, IEEE Trans. Antennas Propogat., 45 (1997), pp. 466–473.
- [8] D. M. BOORE, *Finite-difference methods for seismic wave propagation in heterogeneous materials*, in Methods in Computational Physics, B. Bolt, ed., Academic Press, New York, 1972.
- [9] S. BOYD AND L. VANDENBERGHE, *Convex optimization*, Cambridge university press, 2004.

- [10] J. CARCIONE, *The wave equation in generalized coordinates*, Geophysics, 59 (1994), pp. 1911–1919.
- [11] W. C. CHEW AND W. H. WEEDON, *A 3d perfectly matched medium from modified maxwell's equations with stretched coordinates*, Microwave and Optical Technology Letters, 7 (1994), pp. 599–604.
- [12] D. COLTON AND R. KRESS, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, New York, 2nd ed., 2013.
- [13] P. L. COMBETTES AND J.-C. PESQUET, *Proximal splitting methods in signal processing*, in Fixed-point algorithms for inverse problems in science and engineering, Springer, 2011, pp. 185–212.
- [14] C. W. CONNOR, G. T. CLEMENT, AND K. HYNYNEN, *A unified model for the speed of sound in cranial bone based on genetic algorithm optimization*, Physics in medicine and biology, 47 (2002), pp. 3925–3944.
- [15] S. K. P. D. FINCH AND RAKESH, *Determining a function from its mean values over a family of spheres*, Society of Industrial and Applied Mathematics, 35 (2004), pp. 1213–1240.
- [16] A. T. DE HOOP AND J. H. M. T. VAN DER HIJDEN, *Generation of acoustic waves by an impulsive point source in a fluid/solid configuration with a plane boundary*, J. Acoust. Soc. Am., 75 (1984), pp. 1709–1715, <http://dx.doi.org/10.1121/1.390970>.
- [17] A. T. DEHOOP, *A modification of cagniard's method for solving seismic pulse problems*, Applied Scientific Research, Section B, 8 (1960), pp. 349–356, <http://dx.doi.org/10.1007/BF2920068>.
- [18] A. T. DEHOOP AND J. H. M. T. VAN DER HIJDEN, *Seismic waves generated by an impulsive point source in a solid/fluid configuration with a plane boundary*, Geophysics, 50 (1985), pp. 1083–1090.
- [19] K. FIROUZI, B. COX, B. TREEBY, AND N. SAFFARI, *A first-order k-space model for elastic wave propagation in heterogeneous media*, The Journal of the Acoustical Society of America, 129 (2011), pp. 2611–2611.
- [20] F. J. FRY, *Acoustical properties of the human skull*, J. Acoust. Soc. Am., 63 (1978), pp. 1576–1590.
- [21] P. HELNWEIN, *Some remarks on the compressed matrix representation of symmetric second-order and fourth-order tensors*, Computer Methods in Applied Mechanics and Engineering, 190 (2001), pp. 2753–2770.
- [22] C. HUANG, L. NIE, R. W. SCHOONOVER, Z. GUO, C. O. SCHIRRA, M. A. ANASTASIO, AND L. V. WANG, *Aberration correction for transcranial photoacoustic tomography of primates employing adjunct image data*, Journal of Biomedical Optics, 17 (2012), pp. 066016–1–066016–8, <http://dx.doi.org/10.1117/1.JBO.17.6.066016>.
- [23] C. HUANG, L. NIE, L. V. WANG, AND M. A. ANASTASIO, *Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media*, IEEE. Trans. Med. Imag., 32 (2013), pp. 1097–1110.
- [24] X. JIN, C. LI, AND L. WANG, *Effects of acoustic heterogeneities on transcranial brain imaging with microwave-induced thermoacoustic tomography*, Med. Phys., 35 (2008), pp. 3205–3214, <http://dx.doi.org/10.1118/1.2938731>.
- [25] R. M. JONES, M. A. O'REILLY, AND K. HYNYNEN, *Transcranial passive acoustic mapping with hemispherical sparse arrays using ct-based skull-specific aberration corrections: a simulation study*, Physics in medicine and biology, 58 (2013), p. 4981.
- [26] D. KOMATITSCH AND R. MARTIN, *An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation*, Geophysics, 72 (2007), pp. 155–167.
- [27] R. KRUGER, D. REINECE, AND G. KRUGER, *Thermoacoustic computed tomography- technical considerations*, Med. Phys., 26 (1999), pp. 1832–1837.
- [28] L. A. KUNYANSKY, *Explicit inversion formulae for the spherical mean radon transform*, Inverse Problems, 23 (2007), pp. 373–383.
- [29] M. KUZUOGLU AND R. MITTRA, *Frequency dependence of the constitutive parameters of casual perfectly matched anisotropic absorbers*, IEEE. Microwave Guided Wave Lett., 6 (1996), pp. 447–449.
- [30] C. LI, A. AGUIRRE, J. GAMELIN, A. MAURUDIS, Q. ZHU, AND L. V. WANG, *Real-time photoacoustic tomography of cortical hemodynamics in small animals*, J. Biomed. Opt., 15 (2010), p. 010509, <http://dx.doi.org/10.1117/1.3302807>.
- [31] R. J. LUEBBERS, *Fdtd for nth-order dispersive media*, IEEE. Trans. Antennas Propog., 40 (1992), pp. 1297–1301.
- [32] R. MADARIAGA, K. OLSEN, AND R. ARCHULETA, *Modelling dynamic rupture in a 3D earthquake fault model*, Bulletin of the Seismological Society of America, 88 (1998), pp. 1182–1197.

- [33] F. MARQUET, M. PERNOT, J. AUBRY, G. MONTALDO, L. MARSAC, M. TANTER, AND M. FINK, *Non-invasive transcranial ultrasound therapy based on a 3d ct scan: protocol validation and in vitro results*, Physics in medicine and biology, 54 (2009), p. 2597.
- [34] F. MARQUET, M. PERNOT, J. AUBRY, G. MONTALDO, L. MARSAC, M. TANTER, AND M. FINK, *Non-invasive transcranial ultrasound therapy based on a 3d ct scan: protocol validation and in vitro results*, Physics in medicine and biology, 54 (2009), p. 2597.
- [35] D. MICHÉA AND D. KOMATITSCH, *Accelerating a three-dimensional finite-difference wave propagation code using gpu graphics cards*, Geophysical Journal International, 40 (2010), pp. 389–402.
- [36] P. MICIKEVICIUS, *3d finite difference computation on gpus using cuda*, in Proceedings of 2nd Workshop on General Purpose Processing on Graphics Processing Units, GPGPU-2, New York, NY, USA, 2009, ACM, pp. 79–84, <http://dx.doi.org/10.1145/1513895.1513905>.
- [37] P. MOCZO, J. KRISTEK, M. GALIS, P. PAZAK, AND M. BALAZOVJECH, *The finite-difference and finite-element modeling of seismic wave propagation and earthquake motion*, acta physica slovacica, 57 (2007), pp. 177–406.
- [38] P. MOCZO, J. ROBERTSSON, AND L. EISNER, *The finite-difference time-domain method for modeling of seismic wave propagation*, in Advances in Wave Propagation in Heterogenous Earth, R.-S. Wu and V. Maupin, eds., Elsevier- Academic Press, Cambridge, MA, R. Dmowska ed., 2007, pp. 421–516.
- [39] L. NIE, Z. GUO, AND L. V. WANG, *Photoacoustic tomography of monkey brain using virtual point ultrasonic transducers*, J. Biomed. Opt., 16 (2011), p. 076005, <http://dx.doi.org/10.1117/1.3595842>.
- [40] A. ORAEVSKY AND A. A. KARABUTOV, *Optoacoustic tomography*, in Biomedical Photonics Handbook, T. Vo-Dinh, ed., CRC, Boca Raton, FL, 2003.
- [41] G. PINTON, J.-F. AUBRY, E. BOSSY, M. MUELLER, M. PERNOT, AND M. TANTER, *Attenuation, scattering and absorption of ultrasound in the skull bone*, Medical Physics, 39 (2011), pp. 299–307.
- [42] R. KRUGER, P. LIU, Y. FANG, AND C. R. APPLIEDORN, *Photoacoustic ultrasound (paus) reconstruction tomography*, Med. Phys., 22 (1995), pp. 1605–1609.
- [43] J. A. RODEN AND S. D. GEDNEY, *Convolutional pml (cpml): An efficient fdtd implementation of the cfs-pml for arbitrary media*, Microwave and optical technology letters, 27 (2000), pp. 334–339.
- [44] R. W. SCHOONOVER AND M. A. ANASTASIO, *Compensation of shear waves in photoacoustic tomography with layered acoustic media*, JOSA A, 28 (2011), pp. 2091–2099.
- [45] R. W. SCHOONOVER, L. V. WANG, AND M. A. ANASTASIO, *Numerical investigation of the effects of shear waves in transcranial photoacoustic tomography with a planar geometry*, Journal of Biomedical Optics, 17 (2012), pp. 061215–1–061215–11, <http://dx.doi.org/10.1117/1.JBO.17.6.061215>, <http://dx.doi.org/10.1117/1.JBO.17.6.061215>.
- [46] E. TESSMER AND D. KOSLOFF, *3-D elastic modelling with surface topography by a chebyshev spectral method*, Geophysics, 59 (1994), pp. 464–473.
- [47] B. E. TREEBY, J. JAROS, D. ROHRBACH, AND B. COX, *Modelling elastic wave propagation using the k-wave matlab toolbox*, in 2014 IEEE International Ultrasonics Symposium, IEEE, 2014, pp. 146–149.
- [48] J. VIRIEUX, *P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method*, Geophysics, 51 (1986), pp. 889–901.
- [49] J. VIRIEUX, *SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method*, Geophysics, 49 (1986), pp. 1933–1957.
- [50] G. X. WANG, Y. PANG, AND L. V. WANG, *Non-invasive laser induced photoacoustic tomography for structural and functional in vivo imaging of the brain*, Nat. Biotechnology, 21 (2003), pp. 803–806.
- [51] P. WHITE, G. CLEMENT, AND K. HYNENEN, *Longitudinal and shear mode ultrasound propagation in human skull bone*, Ultrasound in medicine & biology, 32 (2006), pp. 1085–1096.
- [52] M. XU AND L. V. WANG, *Universal back-projection algorithm for photoacoustic computed tomography*, Physical Review E, 71 (2005), p. 016706.
- [53] M. XU AND L. V. WANG, *Photoacoustic imaging in biomedicine*, Rev. Sci. Instrum., 77 (2006), p. 041101.
- [54] Y. XU AND L. V. WANG, *Rhesus monkey brain imaging through intact skull with thermoacoustic tomography*, IEEE. Trans. on Ultrasonics, Ferroelectrics, and Freq. Control, 53 (2006), pp. 542–548, <http://dx.doi.org/10.1109/TUFFC.2006.1610562>.
- [55] Z. XU, Q. ZHU, AND L. V. WANG, *In vivo photoacoustic tomography of mouse cerebral edema induced by cold injury*, J. Biomed. Opt., 16 (2011), p. 066020, <http://dx.doi.org/10.1117/1.3584847>.
- [56] X. YANG AND L. V. WANG, *Monkey brain cortex imaging by photoacoustic tomography*, J. Biomed. Opt.,

13 (2008), p. 044009, <http://dx.doi.org/10.1117/1.2967907>.